Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/EP2003/011352

International filing date: 14 October 2003 (14.10.2003)

Document type: Certified copy of priority document

Document details: Country/Office: EP
Number: 02023560.2
Filing date: 22 October 2002 (22.10.2002)

Date of receipt at the International Bureau: 22 December 2006 (22.12.2006)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

EP/03/11352

Europäisches Patentamt
European Patent Office
Office européen des brevets

Bescheinigung Certificate Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

02023560.2

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

R C van Dijk

Anmeldung Nr:
Application no.: 02023560.2
Demande no:

Anmeldetag:
Date of filing: 22.10.02
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

DeveloGen Aktiengesellschaft für
entwicklungsbiologische Forschung
Rudolf-Wissell-Strasse 28
37079 Göttingen
ALLEMAGNE

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se referer à la description.)

CG14816, CG30346, guf, Mekk1, tws or PP2A-B' homologous proteins involved in the regulation of energy homeostasis

In Anspruch genommene Priorität(en) / Priority(ies) claimed /Priorité(s) revendiquée(s)
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

A61K38/00

Am Anmeldetag benannte Vertragstaaten/Contracting states designated at date of filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR IE IT LI LU MC NL PT SE SK TR

22. Okt. 2002

WEICKMANN & WEICKMANN

Patentanwälte
European Patent Attorneys · European Trademark Attorneys

EPO - Munich
29
22. Okt. 2002

DIPL.-ING. H. WEICKMANN (bis 31.1.01)
DIPL.-ING. F. A. WEICKMANN
DIPL.-CHEM. B. HUBER
DR.-ING. H. LISKA
DIPL.-PHYS. DR. J. PRECHTEL
DIPL.-CHEM. DR. B. BÖHM
DIPL.-CHEM. DR. W. WEISS
DIPL.-PHYS. DR. J. TIESMEYER
DIPL.-PHYS. DR. M. HERZOG
DIPL.-PHYS. B. RUTTENSPERGER
DIPL.-PHYS. DR.-ING. V. JORDAN
DIPL.-CHEM. DR. M. DEY
DIPL.-FORSTW. DR. J. LACHNIT

Unser Zeichen:
29200P EP/WWld

Anmelder:
DeveloGen Aktiengesellschaft für
entwicklungsbiologische Forschung
Rudolf-Wissell-Straße 28

37079 Göttingen

CG14816, CG30346, guf, Mekk1, tws or PP2A-B' homologous proteins involved in the regulation of energy homeostasis

---

Postfach 860 820, 81635 München, Deutschland, Tel. (089) 45563 0, Fax (089) 45563 999, email@weickmann.de

CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologous proteins involved in the regulation of energy homeostasis

Description

This invention relates to the use of nucleic acid sequences encoding CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologous proteins, and the polypeptides encoded thereby and to the use thereof in the diagnosis, study, prevention, and treatment of diseases and disorders related to body-weight regulation, for example, but not limited to, metabolic diseases or dysfunctions such as obesity as well as related disorders such as eating disorder, cachexia, diabetes mellitus, hypertension, coronary heart disease, hypercholesterolemia, dyslipidemia, osteoarthritis, gallstones, cancer, e.g. cancers of the reproductive organs, and sleep apnea.

There are several metabolic diseases of human and animal metabolism, eg., obesity and severe weight loss, that relate to energy imbalance where caloric intake versus energy expenditure is imbalanced. Obesity is one of the most prevalent metabolic disorders in the world. It is still a poorly understood human disease that becomes as a major health problem more and more relevant for western society. Obesity is defined as a body weight more than 20% in excess of the ideal body weight frequently resulting in a significant impairment of health. It is associated with an increased risk for cardiovascular disease, hypertension, diabetes, hyperlipidaemia and an increased mortality rate. Besides severe risks of illness, individuals suffering from obesity are often isolated socially.

Obesity is influenced by genetic, metabolic, biochemical, psychological, and behavioral factors. As such, it is a complex disorder that must be addressed on several fronts to achieve lasting positive clinical outcome.

Since obesity is not to be considered as a single disorder but as a heterogeneous group of conditions with (potential) multiple causes, it is also characterized by elevated fasting plasma insulin and an exaggerated insulin response to oral glucose intake (Koltermann, J. Clin. Invest 65, 1980, 1272-1284). A clear involvement of obesity in type 2 diabetes mellitus can be confirmed (Kopelman, Nature 404, 2000, 635-643).

The molecular factors regulating food intake and body weight balance are incompletely understood. Even if several candidate genes have been described which are supposed to influence the homeostatic system(s) that regulate body mass/weight, like leptin, VCPI, VCPL or the peroxisome proliferator-activated receptor-gamma co-activator, the distinct molecular mechanisms and/or molecules influencing obesity or body weight/body mass regulations are not known. In addition, several single-gene mutations resulting in obesity have been described in mice, implicating genetic factors in the etiology of obesity. (Friedman and Leibel, 1990, Cell 69: 217-220). In the ob mouse a single gene mutation (obese) results in profound obesity, which is accompanied by diabetes (Friedman et. al., 1991, Genomics 11: 1054-1062).

Therefore, the technical problem underlying the present invention was to provide for means and methods for modulating (pathological) metabolic conditions influencing body-weight regulation and/or energy homeostatic circuits. The solution to said technical problem is achieved by providing the embodiments characterized in the claims.

Accordingly, the present invention relates to genes with novel functions in body-weight regulation, energy homeostasis, metabolism, and obesity. The present invention discloses specific genes involved in the regulation of body-weight, energy homeostasis, metabolism, and obesity, and thus in disorders related thereto such as eating disorder, cachexia, diabetes mellitus, hypertension, coronary heart disease, hypercholesterolemia,

dyslipidemia, osteoarthritis, gallstones, cancer, e.g. cancers of the reproductive organs, and sleep apnea. In particular, the present invention describes the human CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologous genes as being involved in those conditions mentioned above.

The term GenBank Accession number relates to NCBI GenBank database entries (Benson et al, Nucleic Acids Res. 28, 2000, 15-18).

The Drosophila CG14816 and its homologous human and mouse gene products might encode a phosphoglycerate mutase (PGAM) enzyme that catalyzes reactions involving the transfer of phospho groups between the three carbon atoms of phosphoglycerate. In mammals, PGAM comprises M-, B- and MB-type isozymes composed of the combination of the muscle-specific (M) and nonmuscle-specific (B) subunits (Uchida K. et al., (1995) Nippon Rinsho 53(5):1247-1252). The catalytic mechanism of PGAM involves the formation of a phosphohistidine intermediate (Fothergill-Gilmore L. A. and Watson H. C., (1989) Adv Enzymol Relat Areas Mol Biol 62:227-313). Phosphoglycerate mutase is an enzyme of the glycolysis and gluconeogenesis that catalyzes the reaction 3-phosphoglycerate into 2-phosphoglycerate and vice versa. The active catalytic core consists of a phosphorylated histidine residue. The enzyme needs a certain amount of 2,3-diphosphoglycerate as a primer for the reaction thereby transfering the 2- or 3-phosphoryl group on the histidine in the active centre.

ATP-NAD kinase catalyzes the phosphorylation of NAD to NADP utilizing ATP and other nucleoside triphosphates as well as inorganic polyphosphate as a source of phosphorus. NADP is essential for biosynthetic pathways, energy, and signal transduction. Its synthesis is catalyzed by NAD kinase. The catalytically active homotetramer is highly selective for its substrates, NAD and ATP. (Lerner F. et al., (2001) Biochem Biophys Res Commun 288(1):69-74). Enzyme activities related to fatty acid synthesis were

determined in liver extracts of rats treated with thioacetamide (TAM). Lipogenesis and cholesterogenesis in vivo were evaluated both in liver and in epididymal adipose tissue. The enzymatic activities of ATP-citrate lyase, acetyl CoA carboxylase, fatty acid synthetase, glycerol kinase and NAD-kinase decrease progressively when TAM was chronically administered (Martin-Sanz P. et al. (1989) Carcinogenesis 10(3):477-481).

The Drosophila gene guf encodes an antizyme that is apparently regulated by translational frameshifting. The mammalian homolog of guf, ornithine decarboxylase antizyme (ODC-AZ) binds to, and destabilizes, ornithine decarboxylase (ODC), a key enzyme in polyamine synthesis. ODC is then rapidly degraded. The expression of ODC-AZ requires programmed, ribosomal frameshifting which is modulated according to the cellular concentration of polyamines. High levels of polyamines induce a +1 ribosomal frameshift in the translation of mRNA for the antizyme leading to the expression of a full- length protein. At least two forms of ODC-AZ exist in mammals. Ornithine decarboxylase (ODC) is among the small set of proteasome substrates that is not ubiquitinated. It is instead degraded in conjunction with the protein antizyme (AZ). ODC and AZ are participants in a regulatory circuit that restricts pools of polyamines, the downstream products of ODC enzymatic activity (Coffino P., (2001) Biochimie 83(3-4):319-323). Antizyme plays an important regulatory role in the synthesis of ornithine decarboxylase (ODC), a key enzyme of polyamine synthesis in higher animals. As well as inactivating polyamine uptake, antizyme is induced by polyamine-enhanced translational frameshifting, and binds to ODC, accelerating its ATP-dependent degradation, a process catalysed by the 26S proteasome (Hayashi S. et al., (1996) Trends Biochem Sci. 21(1):27-30).

A Drosophila MAPK kinase kinase, D-MEKK1, mediates stress responses through activation of p38 MAPK. D-MEKK1 kinase activity was activated in animals under conditions of high osmolarity. Drosophila mutants lacking

D-MEKK1 were hypersensitive to environmental stresses, including elevated temperature and increased osmolarity. In these D-MEKK1 mutants, activation of Drosophila p38 MAPK in response to stress was poor compared with activation in wild-type animals. These results suggest that D-MEKK1 regulation of the p38 MAPK pathway is critical for the response to environmental stresses in Drosophila. D-MEKK1 is structurally similar to the mammalian MEKK4/MTK1 MAPKKK (Inoue H. et al. (2001) EMBO J 20(19):5421-5430). MTK1 (also known as MEKK4) mediates activation of both stress-responsive mitogen-activated protein (MAP) kinases p38 and c-Jun N-terminal kinase (JNK). The MAPK kinase MKK6 is the substrate of MTK1. Growth arrest and DNA damage-inducible 45 (GADD45) proteins were identified as MTK1 activators. GADD45 proteins bind a site in MTK1 near the inhibitory domain and relieve autoinhibition (Mita H. et al. (2002) Mol Cell Biol 22(13):4544-4555). Interleukin-12 and interleukin-18 might activate interferon gamma through GADD45 beta, which activates MEKK4 (Yang J. et al. (2001) Nat Immunol 2(2):157-164).

The tws gene product is a regulatory subunit of protein phosphatase 2 (PP2A), one of the major classes of serine/threonine phosphatases (Uemura T. et al., (1993) 1: Genes Dev. 7(3):429-440). tws is involved in the control of cell division (Omel'ianchuk L. V. et al., (1997) Genetika. 33(11):1494-1501). The beta-catalytic subunit of protein phosphatase 2A (PP2A) is a member of the bestrophin complex. Bestrophin is in the signal transduction pathway that it is regulated by phosphorylation, and that phosphorylation of bestrophin is regulated by PP2A (Marmorstein L. Y. et al., (2002) J Biol Chem 277(34):30591-30597). Type 2A protein phosphatase (PP2A) comprises a diverse family of phosphoserine- and phosphothreonine-specific enzymes ubiquitously expressed in eukaryotic cells. Common to all forms of PP2A is a catalytic subunit (PP2Ac) which can form two distinct complexes, one with a structural subunit termed PR65/A and another with an alpha4 protein. The PR65/A-PP2Ac dimer may further associate with a regulatory subunit and form a trimeric holoenzyme.

To date, three distinct families of regulatory subunits, which control substrate selectivity and phosphatase activity and target PP2A holoenzymes to their substrates, have been identified. Other molecular mechanisms that regulate PP2Ac function include phosphorylation, carboxyl methylation, inhibition by intracellular protein inhibitors (I(1)(PP2A) and I(2)(PP2A)), and stimulation by ceramide. PP2A dephosphorylates many proteins in vitro, but in vivo protein kinases and transcription factors appear to represent two major sets of substrates. (Zolnierowicz S., (2000) Biochem Pharmacol. 60(8):1225-1235). Protein phosphatase 2A (PP2A) is a multimeric enzyme, containing a catalytic subunit complexed with two regulatory subunits. PP2A exerts a range of cellular functions including cell cycle regulation and cell fate determination. (Gotz J. et al., (1998) Proc Natl Acad Sci USA 95:12370-12375).

The Drosophila homolog of the regulatory subunit (B'/PR61) of serine-threonine protein phosphatase 2A (dPP2A-B') specifically interacts with sex combs reduced (SCR) homeodomain. PP2A-B' has an essential role in positively modulating SCR function. (Berry M. and Gehring W., (2000) EMBO J. 19(12):2946-2957). PP2A has been implicated in a variety of regulatory processes including cell growth and division, muscle contraction, and gene transcription (Janssens V. and Goris J., (2001) Biochem J. 353(Pt 3):417-439). PP2A is composed of a 36-kD catalytic subunit C (PP2AC), a highly homologous 65-kD structural subunit (PR65), and any of several different regulatory subunits (B, B', B'' B''') that control its specificity. The B56 family (PR61; B') of regulatory subunits is encoded by at least 5 homologous but distinct genes, termed B56-alpha, -beta, -gamma, -delta, and epsilon (McCright B. et al., (1996) J Biol Chem 271(36):22081-22089). At least three different splice variants exist of the human regulatory subunits B56-γ (γ1, γ2, γ3). B56-γ1, -γ2, -γ3 are localized in the nucleus, B56-δ is found in both the nucleus and the cytoplasm. B56-γ1, -γ2, -γ3 are widely expressed (abundant in heart and muscle), B56-δ is predominantly expressed in brain (McCright B. and

Virshup D. M., (1995) J Biol Chem. 270(44):26123-26128; Tehrani M. A. et al., (1996) J Biol Chem. 271(9):5164-5170). $PP2A_C$ expression level changes during adipocyte differentiation induced by PPAR-γ (Altiok S. et al., (1997) Genes Dev. 11(15):1987-1998). $PP2A_C$ and all B'(PR61; B56) isoforms interact with APC (scaffolding protein for assembly of β-catenin axin and GSK-3β). B56 may direct PP2A to dephosphorylate specific components of the APC-dependent signaling complex and thereby inhibit Wnt signaling (Seeling J. M. et al., (1999) Science. 283(5410):2089-2091).

So far, it has not been described that the CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' proteins of the invention and homologous proteins are involved in the regulation of energy homeostasis and body-weight regulation and related disorders, and thus, no functions in metabolic diseases and dysfunctions and other diseases as listed above have been discussed.

The function of PP2A-B' in metabolic disorders is further validated by data obtained from an additional screen. For example, an additional screen using Drosophila mutants with modifications of the eye phenotype identified a modification of UCP activity by PP2A-B', thereby leading to an altered mitochondrial activity. These findings suggest the presence of similar activities of these described homologous proteins in humans that provides insight into diagnosis, treatment, and prognosis of metabolic disorders.

In this invention, we demonstrate that the correct gene dose of CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' is essential for maintenance of energy homeostasis. The fly Drosophila melanogaster was used as model organism for the identification of proteins involved in the energy homeostasis. Drosophila melanogaster is one of the most intensively studied organisms in biology and serves as a model system for the investigation of many developmental and cellular processes common to

higher eukaryotes, including humans (see, for example, Adams et al., Science 287: 2185-2195 (2000)). The success of Drosophila melanogaster as a model organism is largely due to the power of forward genetic screens to identify the genes that are involved in a biological process (see, Johnston Nat Rev Genet 3: 176-188 (2002); Rorth, Proc Natl Acad Sci U S A 93: 12418-12422 (1996)). In this invention, we have used a genetic screen to identify, that mutations of CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologous genes cause changes in the body weight which is reflected by a significant change in the triglyceride levels. Triglycerides are the most efficient storage for energy in cells, and are significantly increased in obese patients.

In this invention we refer to CG14816 (phosphoglycerate mutase-like), CG30346 (NAD kinase-like), gut feeling (guf), mitogen-activated protein/extracellular signal-regulated kinase (MAP/ERK) kinase kinase 1 (Mekk1), twins (tws), protein phosphatase 2A regularory B subunit (PP2A-B'), CG14816-like, CG30346-like, guf-like, Mekk1-like, tws-like, and PP2A-B'-like, hypothetical protein MGC5352, hypothetical protein FLJ30596 (similar to Y17G7B.10a.p), ornithine decarboxylase antizyme 1, ornithine decarboxylase antizyme 2, ornithine decarboxylase antizyme 3, ornithine decarboxylase antizyme 4, MAP/ERK kinase kinase 4, and protein phosphatase 2 regulatory subunit B (alpha, beta, gamma, and delta isoforms), which include Drosophila and mammalian, preferably human, homologous polypeptides or proteins or sequences encoding those proteins as proteins of the invention.

The present invention discloses that CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologous proteins are regulating the energy homeostasis and fat metabolism especially the metabolism and storage of triglycerides, and polynucleotides, which identify and encode the proteins disclosed in this invention. The invention also relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and

polynucleotides of the invention. The invention also relates to the use of these polynucleotides, polypeptides and effectors thereof in the diagnosis, study, prevention, and treatment of diseases and disorders, for example, but not limited to, metabolic diseases such as obesity as well as related disorders such as eating disorder, cachexia, diabetes mellitus, hypertension, coronary heart disease, hypercholesterolemia, dyslipidemia, osteoarthritis, gallstones, cancer, e.g. cancers of the reproductive organs, and sleep apnea.

CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologous proteins and nucleic acid molecules coding therefore are obtainable from insect or vertebrate species, e.g. mammals or birds. Particularly preferred are nucleic acids encoding the human CG14816, CG30346, guf, Mekk1, tws, or PP2A-B'homologs (in particular the human hypothetical protein MGC5352, human hypothetical protein FLJ30596 (similar to Y17G7B.10a.p), human ornithine decarboxylase antizyme 1, human ornithine decarboxylase antizyme 2, human ornithine decarboxylase antizyme 3, human ornithine decarboxylase antizyme 4, human MAP/ERK kinase kinase 4, and human protein phosphatase 2 regulatory subunit B (alpha, beta, gamma, and delta isoforms).

The invention particularly relates to a nucleic acid molecule encoding a polypeptide contributing to regulating the energy homeostasis and the metabolism of triglycerides, wherein said nucleic acid molecule comprises

(a) the nucleotide sequence of Drosophila CG14816, CG30346, guf, Mekk1, tws, or PP2A-B', human CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologs, and/or a sequence complementary thereto,

(b) a nucleotide sequence which hybridizes at 50°C in a solution containing 1 x SSC and 0.1% SDS to a sequence of (a),

(c) a sequence corresponding to the sequences of (a) or (b) within the degeneration of the genetic code,

(d) a sequence which encodes a polypeptide which is at least 85%, preferably at least 90%, more preferably at least 95%, more preferably at least 98% and up to 99,6% identical to the amino acid sequences of the CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' protein, preferably of the human CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologs,

(e) a sequence which differs from the nucleic acid molecule of (a) to (d) by mutation and wherein said mutation causes an alteration, deletion, duplication and/or premature stop in the encoded polypeptide or

(f) a partial sequence of any of the nucleotide sequences of (a) to (e) having a length of at least 15 bases, preferably at least 20 bases, more preferably at least 25 bases and most preferably at least 50 bases.

The present invention relates to genes with novel functions in body-weight regulation, energy homeostasis, metabolism, and obesity, fragments of said genes, polypeptides encoded by said genes or fragments thereof, and effectors e.g. antibodies, biologically active nucleic acids, such as antisense molecules or ribozymes, aptamers, peptides or low-molecular weight organic compounds recognizing said polynucleotides or polypeptides.

The invention is based on the finding that CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' and the polynucleotides encoding these, are involved in the regulation of triglyceride storage and therefore energy homeostasis. To find genes with novel functions in energy homeostasis, metabolism, and obesity, a functional genetic screen was performed with the model organism Drosophila melanogaster (Meigen). One resource for screening was a Drosophila melanogaster stock collection of EP-lines. The P-vector of this collection has Gal4-UAS-binding sites fused to a basal promoter that can transcribe adjacent genomic Drosophila sequences upon

binding of Gal4 to UAS-sites (Brand A. H. and Perrimon N., (1993) Development 118:401-415; Rorth P., (1996) Proc Natl Acad Sci U S A 93:12418-12422). This enables the EP-line collection for overexpression of endogenous flanking gene sequences. In addition, without activation of the UAS-sites, integration of the EP-element into the gene is likely to cause a reduction of gene activity, and allows determining its function by evaluating the loss-of-function phenotype.

Triglycerides are the most efficient storage for energy in cells, and obese patients mainly show a significant increase in the content of triglycerides. In order to isolate genes with a function in energy homeostasis, several thousand EP-lines were tested for their triglyceride content after a prolonged feeding period (see Examples for more detail). Lines with significantly changed triglyceride content were selected as positive candidates for further analysis. The change of triglyceride content due to the loss of a gene function suggests gene activities in energy homeostasis in a dose dependent manner that controls the amount of energy stored as triglycerides.

In this invention, the content of triglycerides of a pool of flies with the same genotype after feeding for six days was analyzed using a triglyceride assay. Male flies homozygous for the integration of vectors for Drosophila lines HD-EP(X)11032, HD-EP(2)26046, HD-EP(2)20697, or HD-EP(3)31064, or hèterozygous for the integration of vectors for Drosophila lines HD-EP(3)31733, or HD-EP(3)31721 were analyzed in an assay measuring the triglyceride contents of these flies, illustrated in more detail in the EXAMPLES section. The results of the triglyceride content analysis are shown in FIGURE 1, 3, 5, 7, 9, and 11, respectively.

An additional screen using Drosophila mutants with modifications of the eye phenotype identified a modification of UCP activity by PP2A-B', thereby leading to an altered mitochondrial activity.

Genomic DNA sequences were isolated that are localized to the EP vector (herein HD-EP(X)11032, HD-EP(2)26046, HD-EP(2)20697, HD-EP(3)31064, HD-EP(3)31733, and HD-EP(3)31721) integration. Using those isolated genomic sequences public databases like Berkeley Drosophila Genome Project (GadFly; see also FlyBase (1999) Nucleic Acids Research 27:85-88) were screened thereby identifying the integration site of the vectors, and the corresponding gene, described in more detail in the EXAMPLES section. The molecular organization of the gene is shown in FIGURE 2, 4, 6, 8, 10, and 12, respectively.

The present invention further describes polypeptides comprising the amino acid sequences of the proteins of the invention and homologous proteins. Based upon homology, the proteins of the invention and each homologous protein or peptide may share at least some activity.

The invention also encompasses polynucleotides that encode the proteins of the invention and homologous proteins. Accordingly, any nucleic acid sequence, which encodes the amino acid sequences of the proteins of the invention and homologous proteins, can be used to generate recombinant molecules that express the proteins of the invention and homologous proteins. In a particular embodiment, the invention encompasses a nucleic acid encoding Drosophila CG14816, CG30346, guf, Mekk1, tws, or PP2A-B', or human CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologs; referred to herein as the proteins of the invention. It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of nucleotide sequences encoding the proteins, some bearing minimal homology to the nucleotide sequences of any known and naturally occurring gene, may be produced. The invention contemplates each and every possible variation of nucleotide sequence that can be made by selecting combinations based on possible codon choices.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed nucleotide sequences, and in particular, those of the polynucleotide encoding the proteins of the invention, under various conditions of stringency. Hybridization conditions are based on the melting temperature (Tm) of the nucleic acid binding complex or probe, as taught in Wahl, G. M. and S. L. Berger (1987: Methods Enzymol. 152:399-407) and Kimmel, A. R. (1987; Methods Enzymol. 152:507-511), and may be used at a defined stringency. Preferably, hybridization under stringent conditions means that after washing for 1 h with 1 x SSC and 0.1% SDS at 50°C, preferably at 55°C, more preferably at 62°C and most preferably at 68°C, particularly for 1 h in 0.2 x SSC and 0.1% SDS at 50°C, preferably at 55°C, more preferably at 62°C and most preferably at 68°C, a positive hybridization signal is observed. Altered nucleic acid sequences encoding the proteins which are encompassed by the invention include deletions, insertions or substitutions of different nucleotides resulting in a polynucleotide that encodes the same or a functionally equivalent protein.

The encoded proteins may also contain deletions, insertions or substitutions of amino acid residues, which produce a silent change and result in functionally equivalent proteins. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the biological activity of the protein is retained. Furthermore, the invention relates to peptide fragments of the proteins or derivatives thereof such as cyclic peptides, retro-inverso peptides of peptide mimetics having a length of at least 4, preferably at least 6 and up to 50 amino acids.

Also included within the scope of the present invention are alleles of the genes encoding the proteins of the invention and homologous proteins. As used herein, an 'allele' or 'allelic sequence' is an alternative form of the

gene, which may result from at least one mutation in the nucleic acid sequence. Alleles may result in altered mRNAs or polypeptides whose structures or function may or may not be altered. Any given gene may have none, one or many allelic forms. Common mutational changes, which give rise to alleles, are generally ascribed to natural deletions, additions or substitutions of nucleotides. Each of these types of changes may occur alone or in combination with the others, one or more times in a given sequence.

The nucleic acid sequences encoding the proteins of the invention and homologous proteins may be extended utilizing a partial nucleotide sequence and employing various methods known in the art to detect upstream sequences such as promoters and regulatory elements.

In order to express a biologically active protein, the nucleotide sequences encoding the proteins or functional equivalents, may be inserted into appropriate expression vectors, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods, which are well known to those skilled in the art, may be used to construct expression vectors containing sequences encoding the proteins and the appropriate transcriptional and translational control elements. Regulatory elements include for example an promoter, an initiation codons, a stop codon, a mRNA stability regulatory element, and a polyadenylation signal. Expression of a polynucleotide can be assured by (i) constitutive promoters such as the Cytomegalovirus (CMV) promoter/enhancer region, (ii) tissue specific promoters such as the insulin promoter (see, Soria et al., 2000, Diabetes 49:157), SOX2 gene promoter (see Li et al., 1998, Curr. Biol. 8:971-4), Msi-1 promoter (see Sakakibara et al., 1997, J. Neuroscience 17:8300-8312), alpha-cardia myosin heavy chain promoter or human atrial natriuretic factor promoter (Klug et al., 1996, J. clin. Invest 98:216-24; Wu et al., 1989, J. Biol. Chem. 264:6472-79)or (iii) inducible promoters such as the tetracycline inducible

system. Expression vectors can also contain a selection agent or marker gene that confers antibiotic resistance such as the neomycin, hygromycin or puromycin resistance genes. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described in Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, N.Y. and Ausubel, F.M. et al. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York, N.Y.

In a further embodiment of the invention, natural, modified or recombinant nucleic acid sequences encoding the proteins of the invention and homologous proteins may be ligated to a heterologous sequence to encode a fusion protein.

A variety of expression vector/host systems may be utilized to contain and express sequences encoding the proteins or fusion proteins. These include, but are not limited to, micro-organisms such as bacteria transformed with recombinant bacteriophage, plasmid or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (e.g., baculovirus, adenovirus, adeno-associated virus, lentiverus, retrovirus); plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or PBR322 plasmids); or animal cell systems.

The presence of polynucleotide sequences of the invention in a sample can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or portions or fragments of said polynucleotides. Nucleic acid amplification based assays involve the use of oligonucleotides or oligomers based on the sequences specific for the gene to detect transformants containing DNA or RNA encoding the corresponding protein. As used herein 'oligonucleotides' or 'oligomers' refer to a nucleic acid sequence of

at least about 10 nucleotides and as many as about 60 nucleotides, preferably about 15 to 30 nucleotides, and more preferably about 20-25 nucleotides, which can be used as a probe or amplimer.

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting polynucleotide sequences include oligo-labeling, nick translation, end-labeling of labeled RNA probes, PCR amplification using a labeled nucleotide, or enzymatic synthesis. These procedures may be conducted using a variety of commercially available kits (Pharmacia & Upjohn, (Kalamazoo, Mich.); Promega (Madison Wis.); and U.S. Biochemical Corp., (Cleveland, Ohio).

The presence of proteins of the invention in a sample can be determined by immunological methods or activity measurement. A variety of protocols for detecting and measuring the expression of proteins, using either polyclonal or monoclonal antibodies specific for the protein or reagents for determining protein activity are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on the protein is preferred, but a competitive binding assay may be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; Serological Methods, a Laboratory Manual, APS Press, St Paul, Minn.) and Maddox, D. E. et al. (1983; J. Exp. Med. 158:1211-1216).

Suitable reporter molecules or labels, which may be used, include radionuclides, enzymes, fluorescent, chemiluminescent or chromogenic agents as well as substrates, co-factors, inhibitors, magnetic particles, and the like.

The nucleic acids encoding the proteins of the invention can be used to generate transgenic animal or site specific gene modifications in cell lines. Transgenic animals may be made through homologous recombination, where the normal locus of the genes encoding the proteins of the invention is altered. Alternatively, a nucleic acid construct is randomly integrated into the genome. Vectors for stable integration include plasmids, retrovirusses and other animal virusses, YACs, and the like. The modified cells or animal are useful in the study of the function and regulation of the proteins of the invention. For example, a series of small deletions and/or substitutions may be made in the genes that encode the proteins of the invention to determine the role of particular domains of the protein, functions in pancreatic differentiation, etc.

Specific constructs of interest include anti-sense molecules, which will block the expression of the proteins of the invention, or expression of dominant negative mutations. A detectable marker, such as for example lac-Z, may be introduced in the locus of the genes of the invention, where upregulation of expression of the genes of the invention will result in an easily detected change in phenotype.

One may also provide for expression of the genes of the invention or variants thereof in cells or tissues where it is not normally expressed or at abnormal times of development. In addition, by providing expression of the proteins of the invention in cells in which they are not normally produced, one can induce changes in cell behavior.

DNA constructs for homologous recombination will comprise at least portions of the genes of the invention with the desired genetic modification, and will include regions of homology to the target locus. DNA constructs for random integration need not include regions of homology to mediate recombination. Conveniently, markers for positive and/or negative selection are included. Methods for generating cells having targeted gene

modifications through homologous recombination are known in the art. For embryonic stem (ES) cells, an ES cell line may be employed, or embryonic cells may be obtained freshly from a host, e.g. mouse, rat, guinea pig etc. Such cells are grown on an appropriate fibroblast-feeder layer or grown in presence of leukemia inhibiting factor (LIF).

When ES or embryonic cells or somatic pluripotent stem cells have been transformed, they may be used to produce transgenic animals. After transformation, the cells are plated onto a feeder layer in an appropriate medium. Cells containing the construct may be detected by employing a selective medium. After sufficient time for colonies to grow, they are picked and analyzed for the occurrence of homologous recombination or integration of the construct. Those colonies that are positive may then be used for embryo manipulation and blastocyst injection. Blastocysts are obtained from 4 to 6 week old superovulated females. The ES cells are trypsinized, and the modified cells are injected into the blastocoel of the blastocyst. After injection, the blastocysts are returned to each uterine horn of pseudopregnant females. Females are then allowed to go to term and the resulting offspring screened for the construct. By providing for a different phenotype of the blastocyst and the genetically modified cells, chimeric progeny can be readily detected. The chimeric animals are screened for the presence of the modified gene and males and females having the modification are mated to produce homozygous progeny. If the gene alterations cause lethality at some point in development, tissues or organs can be maintained as allogenic or congenic grafts or transplants, or in vitro culture. The transgenic animals may be any non-human mammal, such as laboratory animal, domestic animals, etc. The transgenic animals may be used in functional studies, drug screening, etc.

Diagnostics and Therapeutics

The data disclosed in this invention show that the nucleic acids and proteins of the invention are useful in diagnostic and therapeutic applications implicated, for example but not limited to, in metabolic disorders such as obesity as well as related disorders such as eating disorder, cachexia, diabetes mellitus, hypertension, coronary heart disease, hypercholesterolemia, dyslipidemia, osteoarthritis, gallstones, cancer, e.g. cancers of the reproductive organs, and sleep apnea. Hence, diagnostic and therapeutic uses for the proteins of the invention nucleic acids and proteins of the invention are, for example but not limited to, the following: (i) protein therapeutic, (ii) small molecule drug target, (iii) antibody target (therapeutic, diagnostic, drug targeting/cytotoxic antibody), (iv) diagnostic and/or prognostic marker, (v) gene therapy (gene delivery/gene ablation), (vi) research tools, and (vii) tissue regeneration in vitro and in vivo (regeneration for all these tissues and cell types composing these tissues and cell types derived from these tissues).

The nucleic acids and proteins of the invention and effectors thereof are useful in diagnostic and therapeutic applications implicated in various applications as described below. For example, but not limited to, cDNAs encoding the proteins of the invention and particularly their human homologues may be useful in gene therapy, and the proteins of the invention and particularly their human homologues may be useful when administered to a subject in need thereof. By way of non-limiting example, the compositions of the present invention will have efficacy for treatment of patients suffering from, for example, but not limited to, in metabolic disorders as described above.

The nucleic acids of the invention or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acids or the proteins are to be assessed. Further antibodies that

bind immunospecifically to the novel substances of the invention may be used in therapeutic or diagnostic methods.

For example, in one aspect, antibodies, which are specific for the proteins of the invention and homologous proteins, may be used directly as an effector, e.g. an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express the protein. The antibodies may be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric single chain, Fab fragments, and fragments produced by a Fab expression library. Neutralising antibodies, (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others, may be immunized by injection with the protein or any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. It is preferred that the peptides, fragments or oligopeptides used to induce antibodies to the protein have an amino acid sequence consisting of at least five amino acids, and more preferably at least 10 amino acids.

Monoclonal antibodies to the proteins may be prepared using any technique that provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (K hler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R. J. et al. Proc. Natl. Acad. Sci. 80:2026-2030; Cole, S. P. et al. (1984) Mol. Cell Biol. 62:109-120).

In addition, techniques developed for the production of 'chimeric antibodies', the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity can be used (Morrison, S. L. et al. (1984) Proc. Natl. Acad. Sci. 81:6851-6855; Neuberger, M. S. et al (1984) Nature 312:604-608; Takeda, S. et al. (1985) Nature 314:452-454). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce single chain antibodies specific for the proteins of the invention and homologous proteins. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries (Burton, D. R. (1991) Proc. Natl. Acad. Sci. 88:11120-3). Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening recombinant immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299).

Antibody fragments which contain specific binding sites for the proteins may also be generated. For example, such fragments include, but are not limited to, the F(ab')2 fragments which can be produced by Pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse, W. D. et al. (1989) Science 254:1275-1281).

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding and immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation

between the protein and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reacive to two non-interfering protein epitopes are preferred, but a competitive binding assay may also be employed (Maddox, supra).

In another embodiment of the invention, the polynucleotides of the invention or fragments thereof or nucleic acid effector molecules such as antisense molecules or ribozymes may be used for therapeutic purposes. In one aspect, antisense molecules may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding the proteins of the invention and homologous proteins. Thus, antisense molecules may be used to modulate protein activity or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligomers or larger fragments, can be designed from various locations along the coding or control regions of sequences encoding the proteins. Expression vectors derived from retroviruses, adenovirus, herpes or vaccinia viruses or from various bacterial plasmids may be used for delivery of nucleotide sequences to the targeted organ, tissue or cell population. Methods, which are well known to those skilled in the art, can be used to construct recombinant vectors, which will express antisense molecules complementary to the polynucleotides of the genes encoding the proteins of the invention and homologous proteins. These techniques are described both in Sambrook et al. (supra) and in Ausubel et al. (supra). Genes encoding the proteins of the invention and homologous proteins can be turned off by transforming a cell or tissue with expression vectors, which express high levels of polynucleotides that encode the proteins of the invention and homologous proteins or fragments thereof. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by

endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector and even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing antisense molecules, e.g. DNA, RNA or PNA, to the control regions of the genes encoding the proteins of the invention and homologous proteins, i.e., the promoters, enhancers, and introns. Oligonucleotides derived from the transcription initiation site, e.g., between positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it cause inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J. E. et al. (1994) In; Huber, B. E. and B. I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, N.Y.). The antisense molecules may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples, which may be used, include engineered hammerhead motif ribozyme molecules that can be specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding the proteins of the invention and homologous proteins. Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing

the cleavage site may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Antisense molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences. Such DNA sequences may be incorporated into a variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize antisense RNA constitutively or inducibly can be introduced into cell lines, cells or tissues. RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule or modifications in the nucleobase, sugar and/or phosphate moieties, e.g. the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of non-traditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection and by liposome injections may be achieved using methods, which are well known in the art. Any of the therapeutic methods

described above may be applied to any suitable subject including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of the nucleic acids and the proteins of the invention and homologous nucleic acids or proteins, antibodies to the proteins of the invention and homologous proteins, mimetics, agonists, antagonists or inhibitors of the proteins of the invention and homologous proteins or nucleic acids. The compositions may be administered alone or in combination with at least one other agent, such as stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone or in combination with other agents, drugs or hormones. The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries, which facilitate processing of the active compounds into preparations, which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, Pa.).

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art. For any compounds, the therapeutically effective does can be estimated initially either in cell culture assays, e.g., of preadipocyte cell lines or in animal models, usually mice, rabbits, dogs or pigs. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans. A therapeutically effective dose refers to that amount of active ingredient, for example the nucleid acids or the proteins of the invention and homologous proteins or nucleic acids or fragments thereof, antibodies of the proteins of the invention and homologous proteins, which is sufficient for treating a specific condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between therapeutic and toxic effects is the therapeutic index, and it can be expressed as the ratio, LD50/ED50. Pharmaceutical compositions, which exhibit large therapeutic indices, are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage varies within this range depending upon the dosage from employed, sensitivity of the patient, and the route of administration. The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors, which may be taken into account, include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency

of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week or once every two weeks depending on half-life and clearance rate of the particular formulation. Normal dosage amounts may vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

In another embodiment, antibodies which specifically bind to the proteins may be used for the diagnosis of conditions or diseases characterized by or associated with over- or underexpression of the proteins of the invention and homologous proteins or in assays to monitor patients being treated with the proteins of the invention and homologous proteins, or effectors thereof, e.g. agonists, antagonists, or inhibitors. Diagnostic assays include methods which utilize the antibody and a label to detect the protein in human body fluids or extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by joining them, either covalently or non-covalently, with a reporter molecule. A wide variety of reporter molecules which are known in the art may be used several of which are described above.

A variety of protocols including ELISA, RIA, and FACS for measuring proteins are known in the art and provide a basis for diagnosing altered or abnormal levels of gene expression. Normal or standard values for gene expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibodies to the protein under conditions suitable for complex formation. The amount of standard complex formation may be quantified by various methods, but

preferably by photometry, means. Quantities of protein expressed in control and disease, samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides specific for the proteins of the invention and homologous proteins may be used for diagnostic purposes. The polynucleotides, which may be used, include oligonucleotide sequences, antisense RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which gene expression may be correlated with disease. The diagnostic assay may be used to distinguish between absence, presence, and excess gene expression, and to monitor regulation of protein levels during therapeutic intervention.

In one aspect, hybridization with probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding the proteins of the invention and homologous proteins or closely related molecules, may be used to identify nucleic acid sequences which encode the respective protein. The hybridization probes of the subject invention may be DNA or RNA and derived from the nucleotide sequence of the polynucleotide encoding the proteins of the invention or from a genomic sequence including promoter, enhancer elements, and introns of the naturally occurring gene. Hybridization probes may be labeled by a variety of reporter groups, for example, radionuclides such as ^{32}P or ^{35}S or enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences specific for the proteins of the invention and homologous nucleic acids may be used for the diagnosis of conditions or diseases, which are associated with the expression of the proteins. Examples of such conditions or diseases include, but are not limited to,

pancreatic diseases and disorders, including diabetes. Polynucleotide sequences specific for the proteins of the invention and homologous proteins may also be used to monitor the progress of patients receiving treatment for pancreatic diseases and disorders, including diabetes. The polynucleotide sequences may be used qualitative or quantitative assays, e.g. in Southern or Northern analysis, dot blot or other membrane-based technologies; in PCR technologies; or in dip stick, pin, ELISA or chip assays utilizing fluids or tissues from patient biopsies to detect altered gene expression.

In a particular aspect, the nucleotide sequences specific for the proteins of the invention and homologous nucleic acids may be useful in assays that detect activation or induction of various metabolic diseases such as obesity as well as related disorders such as eating disorder, cachexia, diabetes mellitus, hypertension, coronary heart disease, hypercholesterolemia, dyslipidemia, osteoarthritis, gallstones, cancers of the reproductive organs, and sleep apnea. The nucleotide sequences may be labeled by standard methods, and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the biopsied or extracted sample is significantly altered from that of a comparable have hybridized with nucleotide sequences in the sample, and the presence of altered levels of nucleotide sequences encoding the proteins of the invention and homologous proteins in the sample indicates the presence of the associated disease. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials or in monitoring the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disease associated with expression of the proteins of the invention and homologous proteins, a

normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence or a fragment thereof, which is specific for the nucleic acids encoding the proteins of the invention and homologous nucleic acids, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with those from an experiment where a known amount of a substantially purified polynucleotide is used. Standard values obtained from normal samples may be compared with values obtained from samples from patients who are symptomatic for disease. Deviation between standard and subject values is used to establish the presence of disease. Once disease is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to evaluate whether the level of expression in the patient begins to approximate that, which is observed in the normal patient. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to metabolic diseases such as described above the presence of an unusual amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the metabolic diseases and disorders.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding the proteins of the invention and homologous proteins may involve the use of PCR. Such oligomers may be chemically synthesized, generated enzymatically or produced from a recombinant source. Oligomers will preferably consist of two nucleotide sequences, one

with sense orientation (5'.fwdarw.3') and another with antisense (3'.rarw.5'), employed under optimized conditions for identification of a specific gene or condition. The same two oligomers, nested sets of oligomers or even a degenerate pool of oligomers may be employed under less stringent conditions for detection and/or quantification of closely related DNA or RNA sequences.

In another embodiment of the invention, the nucleic acid sequences may also be used to generate hybridization probes, which are useful for mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome or to a specific region of the chromosome using well known techniques. Such techniques include FISH, FACS or artificial chromosome constructions, such as yeast artificial chromosomes, bacterial artificial chromosomes, bacterial P1 constructions or single chromosome cDNA libraries as reviewed in Price, C. M. (1993) Blood Rev. 7:127-134, and Trask, B. J. (1991) Trends Genet. 7:149-154. FISH (as described in Verma et al. (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York, N.Y.). The results may be correlated with other physical chromosome mapping techniques and genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of the gene encoding the proteins of the invention on a physical chromosomal map and a specific disease or predisposition to a specific disease, may help to delimit the region of DNA associated with that genetic disease.

The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals. In situ hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm

of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms or parts thereof, by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 11q22-23 (Gatti, R. A. et al. (1988) Nature 336:577-580), any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleotide sequences of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier or affected individuals.

In another embodiment of the invention, the proteins of the invention, its catalytic or immunogenic fragments or oligopeptides thereof, an in vitro model, a genetically altered cell or animal, can be used for screening libraries of compounds in any of a variety of drug screening techniques. One can identify effectors, e.g. receptors, ligands or substrates that bind to, modulate or mimic the action of one or more of the proteins of the invention. The protein or fragment thereof employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes, between the proteins of the invention and the agent tested, may be measured. Of particular interest are screening assays for agents that have a low toxicity for mammalian cells. The term "agent" as used herein describes any molecule, e.g. protein or pharmaceutical, with the capability of altering or mimicking the physiological function of one or more of the proteins of the invention. Candidate agents encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 50 and less than about 2,500 Daltons. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or

carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise carbocyclic or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups.

Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, nucleic acids and derivatives, structural analogs or combinations thereof. Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides and oligopeptides. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural or synthetically produced libraries and compounds are readily modified through conventional chemical, physical and biochemical means, and may be used to produce combinatorial libraries. Known pharmacological agents may be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification, amidification, etc. to produce structural analogs. Where the screening assay is a binding assay, one or more of the molecules may be joined to a label, where the label can directly or indirectly provide a detectable signal.

Another technique for drug screening, which may be used, provides for high throughput screening of compounds having suitable binding affinity to the protein of interest as described in published PCT application WO84/03564. In this method, as applied to the proteins of the invention large numbers of different small test compounds, e.g. aptamers, peptides, low-molecular weight compounds etc., are provided or synthesized on a solid substrate, such as plastic pins or some other surface. The test compounds are reacted with the proteins or fragments thereof, and washed. Bound proteins are then detected by methods well known in the

art. Purified proteins can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support. In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding the protein specifically compete with a test compound for binding the protein. In this manner, the antibodies can be used to detect the presence of any peptide, which shares one or more antigenic determinants with the protein.

Finally, the invention also relates to a kit comprising at least one of

(a) a CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' nucleic acid molecule or a fragment thereof;

(b) a vector comprising the nucleic acid of (a);

(c) a host cell comprising the nucleic acid of (a) or the vector of (b);

(d) a polypeptide encoded by the nucleic acid of (a);

(e) a fusion polypeptide encoded by the nucleic acid of (a);

(f) an antibody, an aptamer or another receptor against the nucleic acid of (a) or the polypeptide of (d) or (e) and

(g) an anti-sense oligonucleotide of the nucleic acid of (a).

The kit may be used for diagnostic or therapeutic purposes or for screening applications as described above. The kit may further contain user instructions.

<u>The Figures show:</u>

FIGURE 1 shows the triglyceride content of Drosophila CG14816 (GadFly Accession NumberCG12410) mutants. Shown is the change of triglyceride content of HD-EP(XX)11032 flies caused by integration of the P-vector into

the annotated transcription unit (column 2) in comparison to controls containing all flies of the EP collection ('EP-control', column 1).

FIGURE 2 shows the molecular organization of the mutated CG14816 (Gadfly Accession NumberCG12410) gene locus.

FIGURE 3 shows the triglyceride content of Drosophila CG30346 (GadFly Accession Number) mutants. Shown is the change of triglyceride content of HD-EP(22)26046 flies caused by integration of the P-vector into the annotated transcription unit (column 2) in comparison to controls containing all flies of the EP collection ('EP-control', column 1).

FIGURE 4 shows the molecular organization of the mutated CG30346 (Gadfly Accession Number) gene locus.

FIGURE 5 shows the triglyceride content of Drosophila guf (GadFly Accession Number CG16747) mutants. Shown is the change of triglyceride content of HD-EP(2)20697 flies caused by integration of the P-vector into the annotated transcription unit (column 2) in comparison to controls containing all flies of the EP collection ('EP-control', column 1).

FIGURE 6 shows the molecular organization of the mutated guf (Gadfly Accession Number CG16747) gene locus.

FIGURE 7 shows the triglyceride content of Drosophila Mekk1 (GadFly Accession Number CG7717) mutants. Shown is the change of triglyceride content of HD-EP(3)31064 flies caused by integration of the P-vector into the annotated transcription unit (column 2) in comparison to controls containing 2108 fly lines of the EP collection ('HD-EP collection (N=2108)', column 1).

FIGURE 8 shows the molecular organization of the mutated Mekk1 (Gadfly Accession Number CG7717) gene locus.

FIGURE 9 shows the triglyceride content of Drosophila tws (GadFly Accession Number CG6235) mutants. Shown is the change of triglyceride content of HD-EP(3)31733/TM3,Ser flies caused by heterozygous integration of the P-vector into the annotated transcription unit (column 2) in comparison to controls containing all flies of the EP collection ('EP-control', column 1).

FIGURE 10 shows the molecular organization of the mutated tws (Gadfly Accession Number CG6235) gene locus.

FIGURE 11 shows the triglyceride content of Drosophila PP2A-B' (GadFly Accession Number CG7913) mutants. Shown is the change of triglyceride content of HD-EP(2)26046 / elav flies caused by ectopic expression of the PP2A-B' gene mainly in the neurons of these flies (column 2) in comparison to controls containing all flies of the EP / elav collection ('EP / elav control', column 1).

FIGURE 12 shows the molecular organization of the mutated PP2A-B' (Gadfly Accession Number CG7913) gene locus.

<u>The examples illustrate the invention:</u>

Example 1: Measurement of triglyceride content

Mutant flies are obtained from a fly mutation stock collection. The flies are grown under standard conditions known to those skilled in the art. In the course of the experiment, additional feedings with bakers yeast

(Saccharomyces cerevisiae) are provided for the EP-lines HD-EP(X)11032, HD-EP(2)26046, HD-EP(2)20697, HD-EP(3)31064, HD-EP(3)31733, and HD-EP(3)31721. The average change of triglyceride content of Drosophila containing the EP-vector as homozygous/heterozygous viable integration was investigated in comparison to control flies grown under the same conditions (see FIGURES 1, 3, 5, 7, 9, and 11, respectively). For determination of triglyceride content, flies were incubated for 5 min at 90°C in an aqueous buffer using a waterbath, followed by hot extraction. After another 5 min incubation at 90°C and mild centrifugation, the triglyceride content of the flies extract was determined using Sigma Triglyceride (INT 336-10 or -20) assay by measuring changes in the optical density according to the manufacturer's protocol. As a reference the protein content of the same extract was measured using BIO-RAD DC Protein Assay according to the manufacturer's protocol. These experiments and assays were repeated several times.

The average triglyceride level of all flies of the EP collection (referred to as EP-control) is shown as 100% in the first columns in FIGURES 1, 3, 5, and 9. The average triglyceride level of 2108 flies of the HD-EP collection (referred to as HD-EP collection (N=2108)) is shown as 100% in the first column in FIGURE 7. The average triglyceride level of all flies containing the elav-Gal4 vector (referred to as EP / elav control) is shown as 100% in the first column in FIGURE 11. Standard deviations of the measurements are shown as thin bars.

HD-EP(X)11032 homozygous flies show constantly a lower triglyceride content than the controls (column 2 in FIGURE 1, 'HD-EP11032'). HD-EP(2)26046 homozygous flies show constantly a lower triglyceride content than the controls (column 2 in FIGURE 3, ' HD-EP26046'). HD-EP(2)20697 homozygous flies show constantly a lower triglyceride content than the controls (column 2 in FIGURE 5, 'HD-EP20697'). HD-EP(3)31064 homozygous flies show constantly a higher triglyceride

content than the controls (column 2 in FIGURE 7, 'HD-EP31064'). HD-EP(3)31733 heterozygous flies, containing a TM3,Ser balancer chromosome, show constantly a lower triglyceride content than the controls (column 2 in FIGURE 9, 'HD-EP31733/TM3,Ser'). Therefore, the loss of gene activity is responsible for changes in the metabolism of the energy storage triglycerides. The offspring of HD-EP(3)31721 males that are crossed to elav-Gal4 virgins, carrying a copy of the HD-EP(3)31721 vector and a copy of the elav-Gal4 vector, leading to ectopic expression of adjacent genomic DNA sequences 3prime of the HD-EP(3)31721 integration locus, mainly in neurons of these flies, shows constantly a higher triglyceride content than the controls (column 2 in FIGURE 11, 'HD-EP(3)31721 /elav'). Therefore, the gain of gene activity in the neurons is responsible for changes in the metabolism of the energy storage triglycerides.

Example 2: Identification of the genes

(i) CG14816

Genomic DNA sequences were isolated that are localized directly adjacent to the EP vector (herein HD-EP(X)11032) integration. Using those isolated genomic sequences public databases like Berkeley Drosophila Genome Project (GadFly) were screened thereby confirming the homozygous viable integration site of the HD-EP(X)11032 HD-EP(X)10646 vector in base pair 1 of the of expressed sequence tag (EST) SD17949 that overlaps with the cDNA of CG14816 in sense orientation. FIGURE 2 shows the molecular organization of this gene locus. The chromosomal localization site of integration of the vector of HD-EP(X)11032 is at gene locus X, 2B13 (according to Flybase), 2A2 (according to Gadfly release 2). In FIGURE 2, genomic DNA sequence is represented by the assembly as a dotted grey line in the middle that includes the integration site of the vector for line HD-EP(X)11032. Numbers represent the coordinates of the genomic DNA (starting at position 1619774 on chromosome X, ending at position

1621337 on chromosome X). The insertion site of the P-element in Drosophila line HD-EP(X)11032 is shown as triangle in the "P Elements -" line and is labeled. Dark grey boxes on the "cDNA +" and the "cDNA- " lines, linked by light grey boxes, represent the predicted genes (as predicted by the Berkeley Drosophila Genome Project, GadFly and by Magpie). Predicted exons are shown as dark grey boxes, predicted introns are shown as light grey boxes. The gene CG14816 is labeled (referred to as EG:63B12.4 (CG14816)). Transcribed DNA sequences (ESTs) are shown as grey bars in the "EST -" line. Therefore, expression of the cDNA encoding 2CG14816 could be affected by integration of the vector of line HD-EP(X)11032, leading to a change in the amount of energy storage triglycerides.

(ii) CG30346

Genomic DNA sequences were isolated that are localized directly adjacent to the EP vector (herein HD-EP(2)26046) integration. Using those isolated genomic sequences public databases like Berkeley Drosophila Genome Project (GadFly) were screened thereby confirming the homozygous viable integration site of the HD-EP(2)26046 into the third exon of CG30346-RA and last exon of CG30346-RB in antisense orientation. FIGURE 4 shows the molecular organization of this gene locus. The chromosomal localization site of integration of the vector HD-EP(2)26046 is at gene locus 2R2R, 45A2. In FIGURE 4, genomic DNA sequence is represented by the assembly as a black arrow in the middle of the figure that includes the integration site of HD-EP(2)26046. Ticks represent the length in basepairs of the genomic DNA (1000 base pairs per tick). The insertion site of the P-element in Drosophila line HD-EP(2)26046 is shown as triangle and is labeled. Dark grey bars, linked by dark grey lines represent cDNAs of the predicted genes (as predicted by the Berkeley Drosophila Genome Project, GadFly release 3). Predicted exons of the Drosophila cDNA CG30346 (GadFly Accession Number) are shown as dark grey bars and predicted introns as slim grey lines in the lower half of the figure. Therefore,

expression of the cDNA encoding CG30346 could be affected by integration of vector of line HD-EP(2)26046, leading to a change in the amount of energy storage triglycerides.

(iii) guf

Genomic DNA sequences were isolated that are localized directly adjacent to the EP vector (herein HD-EP(2)20697) integration. Using those isolated genomic sequences public databases like Berkeley Drosophila Genome Project (GadFly) were screened thereby confirming the homozygous viable integration site of the HD-EP(2)20697 105 base pairs 5prime of CG16747-RA in sense orientation. FIGURE 6 shows the molecular organization of this gene locus. The chromosomal localization site of integration of the vector HD-EP(2)20697 is at gene locus 2R, 48E4. In FIGURE 6, genomic DNA sequence is represented by the assembly as a black arrow in the middle of the figure that includes the integration site of HD-EP(2)20697. Ticks represent the length in basepairs of the genomic DNA (1000 base pairs per tick). The insertion site of the P-element in Drosophila line HD-EP(2)20697 is shown as triangle and is labeled. Dark grey bars, linked by dark grey lines represent cDNAs of the predicted genes (as predicted by the Berkeley Drosophila Genome Project, GadFly release 3). Predicted exons of the Drosophila cDNA guf (GadFly Accession Number CG16747) are shown as dark grey bars and predicted introns as slim grey lines in the lower half of the figure. Therefore, expression of the cDNA encoding guf could be affected by integration of vector of line HD-EP(2)20697, leading to a change in the amount of energy storage triglycerides.

(iv) Mekk1

Genomic DNA sequences were isolated that are localized directly adjacent to the EP vector (herein HD-EP(3)31064) integration. Using those isolated genomic sequences public databases like Berkeley Drosophila Genome Project (GadFly) were screened thereby confirming the homozygous viable

integration site of the HD-EP(3)31064 13 base pairs 5prime of Mekk1 in antisense orientation. FIGURE 8 shows the molecular organization of this gene locus. The chromosomal localization site of integration of the vector HD-EP(3)31064 is at gene locus 3R, 91B8-C1 (according to Flybase), 91C5 (according to Gadfly release 3). In FIGURE 8, genomic DNA sequence is represented by the assembly as a black arrow in the middle of the figure that includes the integration site of HD-EP(3)31064. Ticks represent the length in basepairs of the genomic DNA (1000 base pairs per tick). The insertion site of the P-element in Drosophila line HD-EP(3)31064 is shown as triangle and is labeled. Dark grey bars, linked by dark grey lines represent cDNAs of the predicted genes (as predicted by the Berkeley Drosophila Genome Project, GadFly release 3). Predicted exons of the Drosophila cDNA Mekk1 (GadFly Accession Number CG7717) are shown as dark grey bars and predicted introns as slim grey lines in the lower half of the figure. Therefore, expression of the cDNA encoding Mekk1 could be affected by integration of vector of line HD-EP(3)31064, leading to a change in the amount of energy storage triglycerides.

(v) tws

Genomic DNA sequences were isolated that are localized directly adjacent to the EP vector (herein HD-EP(3)31733) integration. Using those isolated genomic sequences public databases like Berkeley Drosophila Genome Project (GadFly) were screened thereby confirming the heterozygous viable integration site of the HD-EP(3)31733 into the second intron of CG6235-RA, and CG6235-RF, and into the second exon of CG6235-RB, CG6235-RC, CG6235-RD, CG6235-RE, CG6235-RG, and CG6235-RH in antisense orientation. FIGURE 10 shows the molecular organization of this gene locus. The chromosomal localization site of integration of the vector HD-EP(3)31733 is at gene locus 3R, 85F10-12 (according to Flybase), 85F13-14 (according to Gadfly release 3). In FIGURE 10, genomic DNA

sequence is represented by the assembly as a black arrow in the upper half of the figure that includes the integration site of HD-EP(3)31733. Ticks represent the length in basepairs of the genomic DNA (1000 base pairs per tick). The insertion site of the P-element in Drosophila line HD-EP(3)31733 is shown as triangle and is labeled. Dark grey bars, linked by dark grey lines represent cDNAs of the predicted genes (as predicted by the Berkeley Drosophila Genome Project, GadFly release 3). Predicted exons of the Drosophila cDNA tws (GadFly Accession Number CG6235) are shown as dark grey bars and predicted introns as slim grey lines in the middle and the lower half of the figure. Therefore, expression of the cDNA encoding tws could be affected by integration of vector of line HD-EP(3)31733, leading to a change in the amount of energy storage triglycerides.

(vi) PP2A-B'

Genomic DNA sequences were isolated that are localized directly adjacent to the EP vector (herein HD-EP(3)31721) integration. Using those isolated genomic sequences public databases like Berkeley Drosophila Genome Project (GadFly) were screened thereby confirming the homozygous viable integration site of the HD-EP(3)31721 into the first intron of PP2A-B' in sense orientation. FIGURE 12 shows the molecular organization of this gene locus. The chromosomal localization site of integration of the vector HD-EP(3)31721 is at gene locus 3R, 90E3-4. In FIGURE 12, genomic DNA sequence is represented by the assembly as a black arrow in the lower half of the figure that includes the integration site of HD-EP(3)31721. Ticks represent the length in basepairs of the genomic DNA (10000 base pairs per tick). The insertion site of the P-element in Drosophila line HD-EP(3)31721 is shown as triangle and is labeled. Dark grey bars, linked by dark grey lines represent cDNAs of the predicted genes (as predicted by the Berkeley Drosophila Genome Project, GadFly release 3). Predicted exons of the Drosophila cDNA tws (GadFly Accession Number CG6235) are shown as dark grey bars and predicted introns as slim grey lines in the middle and the upper half of the figure. Therefore, expression of the cDNA

encoding PP2A-B' could be affected by integration of vector of line HD-EP(3)31721, leading to a change in the amount of energy storage triglycerides.

Table 1 is summarizing the data of our molecular analysis of the Drosophila proteins identified in this invention as being involved in the regulation of the metabolism.

Table 1. Molecular analysis of Drosophila CG14816, CG30346, guf, Mekk1, tws, and PP2A-B'

Analysis	**Genetic interaction**
CG14816	not described (Flybase)
CG30346	not described (Flybase)
guf	not described (Flybase)
Mekk1	not described (Flybase)
tws	Cdc27, shotgun (E-Cadherin) and armadillo (¤-Catenin) (Flybase)
PP2A-B'	with *Scr*
Analysis	**Protein**
CG14816	Phosphoglycerate mutase-like
CG30346	potential NAD kinase
guf	ornithine decarboxylase inhibitor (Flybase)
Mekk1	MAP kinase kinase kinase involved in protein amino acid phosphorylation (Flybase)
tws	protein phosphatase type 2A, regulator nvolved in the regulation of mitosis which is localised to the cytoplasm (Flybase)
PP2A-B'	protein phosphatase type 2A, regulator
Analysis	**Protein domains**
CG14816	Phosphoglycerate mutase-like, Phosphoglycerate mutase family

CG30346	not described (Flybase)
guf	Ornithine decarboxylase antizyme (Flybase)
Mekk1	Eukaryotic protein kinase & Serine/Threonine protein kinase family active site (Flybase)
tws	Protein phosphatase 2A regulatory subunit PR55, Trp-Asp repeat (WD-repeat) (Flybase)
PP2A-B'	Protein phosphatase 2A regulatory B subunit (B56 family)
Analysis	**InterPro analysis**
CG14816	Phosphoglycerate/bisphosphoglycerate mutase (IPR001345)
CG30346	CG30346-RA: ATP-NAD kinase (IPR002504)
guf	Ornithine decarboxylase antizyme (IPR002993)
Mekk1	Eukaryotic protein kinase (IPR000719), Tyrosine protein kinase (IPR001245), Serine/Threonine protein kinase (IPR002290)
tws	Protein phosphatase 2A regulatory subunit PR55 (IPR000009), G-protein beta WD-40 repeat (IPR001680)
PP2A-B'	-
Analysis	**Locus**
CG14816	X, 2B13 (Flybase); X, 2A2 (Gadfly release 2)
CG30346	2R, 45A2 (Flybase); 2R, 45A2 (Gadfly release 3)
guf	2R, 48E4 (Flybase); 2R 48E4 (Gadfly release 3)
Mekk1	3R, 91B8-C1 (Flybase); 3R, 91C5 (Gadfly release 3)
tws	3R, 85F10-12 (Flybase); 3R, 85F13-14 (Gadfly release 3)
PP2A-B'	3R 90E3-4 (GadFly)
Analysis	**Ests**
CG14816	several including GH02880
CG30346	several including GH09647 and RE43986
guf	many including RE14453 and GH26763
Mekk1	AT24611, AT31494, AT14062, LD30033, LD39970, GM28261, LD42182, LD40836, SD09178, SD20835 (Gadfly release 3)
tws	GM31008, AI945059, HL01840, HL01804, SD20892, GH11683, RE60168, GM08394, RE72358, RE73279, RE34913, SD20465, LD43046, AT16930, LD14078, RH41581, RE01343, AT23584, AT04854, AT22386, GH28839, LD03262, LD12394, RE66660, RE53165, RE52128, RE31837, RE04309, RE17579
PP2A-B'	-
Analysis	**CDNA**

CG14816	AA950742 (712 base pairs mRNA), AI064660 (564 base pairs mRNA), AW940112 (493 base pairs mRNA), AY060608 (1104 base pairs mRNA; protein:AAL28156)
CG30346	not described (Flybase)
guf	AA803901 (535 base pairs mRNA), AF038597 (1654 base pairs mRNA; protein:AAC97538), AI513707 (853 base pairs mRNA), AI519013 (575 base pairs, mRNA), AW941028 (538 base pairs mRNA), AY069220 (1733 base pairs mRNA; protein:AAL39365), AY071071 (1755 base pairs mRNA; protein:AAL48693), BI172172 (606 base pairs mRNA), U29529 (2163 base pairs mRNA; protein:AAB49330) (Flybase)
Mekk1	AB069961 (5176 base pairs mRNA; protein:BAB62891), AB069962 (5464 base pairs mRNA; protein:BAB62892), AI542730 (534 base pairs mRNA), AW944582 (474 base pairs mRNA) (Flybase)
tws	AA438587 (651 base pairs mRNA), (597 base pairs mRNA), AI945059 (306 base pairs mRNA), AW942232 (498 base pairs mRNA), AY061152 (2517 base pairs mRNA; protein:AAL28700), BG638058 (570 base pairs mRNA), D13004 (3095 base pairs mRNA; protein:BAA02367), L07581 (2076 base pairs mRNA; protein:AAA99870), L07583 (2899 base pairs mRNA; protein:AAA99871) (Flybase)
PP2A-B'	AA949854 (735 base pairs mRNA), AJ277140 (2250 base pairs mRNA; protein:CAB86364); AW942997 (519 base pairs mRNA)
Analysis	**genomic DNA**
CG14816	AE003422 (300739 base pairs DNA; protein: AAF45678), AL021106 (36401 base pairs DNA; protein: CAA15939) (Flybase)
CG30346	AE003834 (257324 base pairs DNA; protein:AAF58992; protein:AAF58993; protein:AAM68802) (Flybase)
guf	AE003823 (262731 base pairs DNA; protein:AAF58567; protein:AAF58569; protein:AAF58570) (Flybase)
Mekk1	AE003723 (255624 base pairs DNA; protein:AAF55592) (Flybase)
tws	AC009183 (192055 base pairs DNA), AE003685 (225038 base pairs DNA; protein:AAF54498; protein:AAF54499) (Flybase)
PP2A-B'	-
Analysis	**NCBI locus ID**
CG14816	31143, Dm EG:63B12.4, 2B13; Aliases: CG14816, CT34629; RefSeq: NM_130595; Nucleotide: AE003422, AL021106, AA950742,

	AW940112, AY060608; Protein: NP_569951, AAF45678, CAA15939, AAL28156 (all 289 amino acids)
CG30346	246554, Dm CG30346, 45A8; Nucleotide: AE003834; Protein: AAF58992 (367 amino acids), AAF58993 (413 amino acids), AAM68802 (409 amino acids)
guf	36307, Dm guf, gut feeling, 48F1; Aliases: AZ, ODA, CG16747, CT37253, CT41663, CT41665, CT41667, l(2)02833, l(2)131/7, l(2)154/1, l(2)90/37, l(2)k09037, ornithine decarboxylase antizyme; Nucleotide: AE003823, AQ025598, AQ025599, AQ073369, AZ300928, AA803901, AF038597, AI519013, AY069220, AY071071, U29529; Protein: AAF58567 (270 amino acids), AAF58568 (199 amino acids), AAF58569 (254 amino acids), AAF58570 (248 amino acids), AAC97538 (254 amino acids), AAL39365 (119 amino acids), AAL48693 (68 amino acids), AAB49330 (186 amino acids)
Mekk1	42253, Dm Mekk1, 91C5; Aliases: MEKK, MEKK4, CG7717, CT20838, DmMEKK1; RefSeq: NM_142493; Nucleotide: AE003723, AQ254736, BH759341, BH759342, AB069961, AB069962, AI542730, AW944582, AY118659; Protein: NP_650750 (1497 amino acids), AAF55592 (1612 amino acids), BAB62891 (1571 amino acids), BAB62892 (1497 amino acids), AAM50028 (1571 amino acids)
tws	47877, Dm tws, twins, 85F13-85F14; Aliases: tw, aar, 2414, PP2A, PR55, v158, DPR55, PP2Ac, B/PR55, CG6235, CT19500, CT36963, Pp2A-85F, l(3)01436, l(3)02414, l(3)j11C8, l(3)s1801, phosphoprotein phosphatase 2A 55 kDa; RefSeq: NM_057532, NM_057533, NM_134284, NM_134285, NM_134286; Nucleotide: AC009183, AE003685, G00588, G01405, L07585, L07586, L12544, AA438587, AW942232, AY061152, BG638058, D13004, L07581, L07583; Protein: NP_476880, NP_476881, NP_599111, NP_599112, NP_599113, AAF54498, AAF54499, AAB00371, AAB00372, AAL28700, BAA02367, AAA99870, AAA99871
PP2A-B'	42169, Dm *PP2A-B'*; Aliases: CG7901, CG7913, B'/PR61, CT23882, CT23900, CT42545, dPP2A,B'; RefSeq: NM_142424; Nucleotide: AE003721, BH256430, AJ277140, AW942997, AY118546; Protein: NP_650681, AAF55499, AAF55500, CAB86364, AAM49915
Analysis	**Drosophila mutations & mutants**
CG14816	not described (Flybase)

CG30346	not described (Flybase)
guf	There are 14 recorded alleles, 13 classical mutants (2 available from the public stock centers) and 1 wild-type.
Mekk1	There are 3 recorded alleles, 2 classical mutants (1 available from the public stock centers) and 1 wild-type (Flybase)
tws	There are 18 recorded alleles: 1 in vitro construct (not available from the public stock centers), 15 classical mutants (1 available from the public stock centers) and 2 wild-type.
PP2A-B'	Injection of embryos with *PP2A-B'* dsRNA results in a complete loss of salivary glands in about 40% of cases. (Berry and Gehring, 2000)
Analysis	**Phenotypic info**
CG14816	-
CG30346	-
guf	Amorphic mutations have been isolated which affect the embryonic peripheral nervous system, the embryonic central nervous system, the embryonic/larval somatic muscle and the embryonic visceral muscle and are embryonic recessive lethal. (Flybase)
Mekk1	-
tws	Hypomorphic mutations have been isolated which affect the trichogen cell, the tormogen cell, the macrochaeta and other tissues and are adult recessive lethal, recessive behavioural and uncoordinated. tws is an essential gene required during embryogenesis and late larval development. Mutations cause abnormalities in mitosis, both in the early embryo and in the larval brain. Mutant alleles are pupal lethal and cause pattern duplication in imaginal discs. (Flybase)
PP2A-B'	-

Example 3: Identification of human CG14816, CG30346, guf, Mekk1, tws, and PP2A-B' proteins

CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologous proteins and nucleic acid molecules coding therefore are obtainable from insect or

vertebrate species, e.g. mammals or birds. Particularly preferred are nucleic acids comprising Drosophila CG14816, CG30346, guf, Mekk1, tws, and PP2A-B', or human CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologs (in particular the human hypothetical protein MGC5352, human hypothetical protein FLJ30596 (similar to Y17G7B.10a.p), human ornithine decarboxylase antizyme 1, human ornithine decarboxylase antizyme 2, human ornithine decarboxylase antizyme 3, human ornithine decarboxylase antizyme 4, human MAP/ERK kinase kinase 4, and human protein phosphatase 2 regulatory subunit B (alpha, beta, gamma, and delta isoforms).

Sequences homologous to Drosophila CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' were identified using the publicly available program BLASTP 2.2.3 of the non-redundant protein data base of the National Center for Biotechnology Information (NCBI)(see, Altschul et al., 1997, Nucleic Acids Res. 25:3389-3402).

Table 2 shows the best human homologs of the Drosophila CG14816, CG30346, guf, Mekk1, tws, and PP2A-B' protein.

Table 2. Human homolog proteins to Drosophila CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' proteins

II. CG14816
NCBI (National Center for Biotechnology Information) human locus identification (ID): 192111, Hs MGC5352, hypothetical protein MGC5352
RefSeq[R]: GenBank Accession Number NM_138575
Nucleotide: GenBank Accession Number BC008196
Protein: GenBank Accession Numbers NP_612642, AAH08196 (both 255 amino acids)

II. CG 30346

NCBI (National Center for Biotechnology Information) human locus identification (ID): 133686, Hs FLJ30596, hypothetical protein FLJ30596, 5p13.1

RefSeq[R]: GenBank Accession Number NM_153013, XM_059667

Nucleotide: GenBank Accession Number AK055158

Protein: GenBank Accession Numbers NP_694558, XP_059667, BAB70864 (all 279 amino acids)

III. guf

NCBI (National Center for Biotechnology Information) human locus identification (ID): 4946, Hs OAZ1, ornithine decarboxylase antizyme 1, 19p13.3

Aliases: OAZ

OMIM: 601579

RefSeq[R]: GenBank Accession Number NM_004152

Nucleotide: GenBank Accession Numbers AC004152, D89870, BC007247, BC010063, D78361, D87914, U09202

Protein: GenBank Accession Numbers NP_004143 (228 amino acids), AAC02802 (68 amino acids), AAC02803 (160 amino acids), BAA23101 (228 amino acids), BAA11373 (68 amino acids), BAA11374 (191 amino acids), BAA13497 (228 amino acids), AAA82154 (68 amino acids), AAA82155 (226 amino acids)

NCBI (National Center for Biotechnology Information) human locus identification (ID): 4947, Hs OAZ2, ornithine decarboxylase antizyme 2, 15q22.1

OMIM: 604152

RefSeq[R]: GenBank Accession Number NM_002537

Nucleotide: GenBank Accession Numbers AF057297, AF242521

Protein: GenBank Accession Numbers NP_002528, AAD03265, AAF99601 (all 189 amino acids)

NCBI (National Center for Biotechnology Information) human locus identification (ID): 51686, Hs OAZ3, ornithine decarboxylase antizyme 3, 1q21.1

Aliases: OAZ-t

OMIM: 605138

RefSeq[R]: GenBank Accession Number NM_016178

Nucleotide: GenBank Accession Number AF175296

Protein: GenBank Accession Numbers NP_057262, AAD51734 (both 187 amino acids)

NCBI (National Center for Biotechnology Information) human locus identification (ID): ornithine decarboxylase antizyme 4,

Nucleotide: GenBank Accession Number AF293339

Protein: GenBank Accession Number AAG26885 (218 amino acids)

IV. Mekk1

NCBI (National Center for Biotechnology Information) human locus identification (ID): 4216, Hs MAP3K4, mitogen-activated protein kinase kinase kinase 4, 6q25.3

Aliases: MTK1, MEKK4, MAPKKK4, KIAA0213

OMIM: 602425

RefSeq[R]: GenBank Accession Numbers NM_005922, NM_006724

Nucleotide: GenBank Accession Numbers AF002715, D86968

Protein: GenBank Accession Numbers NP_005913 (1607 amino acids), NP_006715 (1558 amino acids), AAB68804 (1607 amino acids), BAA13204 (1491 amino acids)

V. tws

NCBI (National Center for Biotechnology Information) human locus identification (ID): 5520, Hs PPP2R2A, protein phosphatase 2 (formerly 2A), regulatory subunit B, 8p21.1 (PR 52), alpha isoform

OMIM: 604941

RefSeq: GenBank Accession Number NM_002717

Nucleotide: GenBank Accession Number M64929

Protein: GenBank Accession Numbers NP_002708, AAA36490 (both 447 amino acids)

NCBI (National Center for Biotechnology Information) human locus identification (ID): 55844, Hs MDS026, uncharacterized hematopoietic stem/progenitor cells protein MDS026, 10

RefSeq: GenBank Accession Number NM_018461

Nucleotide: GenBank Accession Numbers AB040974, AF220046, AK001483

Protein: GenBank Accession Numbers NP_060931 (137 amino acids), BAA96065 (477 amino acids), AAF67639 (137 amino acids)

NCBI (National Center for Biotechnology Information) human locus identification (ID): 5521, Hs PPP2R2B, protein phosphatase 2 (formerly 2A), regulatory subunit B (PR 52), beta isoform, 5q31-5q33

Aliases: SCA12, PR55-BETA

OMIM: 604325

RefSeq: GenBank Accession Number NM_004576

Nucleotide: GenBank Accession Number M64930

Protein: GenBank Accession Number NP_004567, AAA36493 (both 443 amino acids)

NCBI (National Center for Biotechnology Information) human locus identification (ID): 5522, Hs PPP2R2C, protein phosphatase 2 (formerly 2A), regulatory subunit B (PR 52), gamma isoform, 4p16

Aliases: PR52

OMIM: 605997

RefSeq: GenBank Accession Number XM_029744

Nucleotide: GenBank Accession Numbers AF096160, AF086924

Protein: GenBank Accession Numbers XP_029744 (292 amino acids), AAD20987 (423 amino acids), AAG39636 (447 amino acids)

VI. PP2A-B'

NCBI (National Center for Biotechnology Information) human locus identification (ID): 5527, Hs PPP2R5C, protein phosphatase 2, regulatory subunit B (B56), gamma isoform, 3p21

Aliases: B56G

OMIM: 601645

RefSeq [R]: GenBank Accession Number NM_002719

Nucleotide: GenBank Accession Numbers AY052369, BC016183, D26445, L42375, U37352, Z69030

Protein: GenBank Accession Numbers NP_002710, NP_002710, AAL14777, AAH16183, BAA05465, AAC37603, AAC50387, CAA93154

NCBI (National Center for Biotechnology Information) human locus identification (ID): 5528, Hs PPP2R5D, protein phosphatase 2, regulatory subunit B (B56), delta isoform, 6p21.1

OMIM: 601646

RefSeq: GenBank Accession Number NM_006245

Nucleotide: GenBank Accession Numbers AB000634, BC001095, BC001175, BC010692, L76702

Protein: GenBank Accession Numbers NP_006236, NP_006236, BAA20381, AAH01095, AAH01175, AAH10692, AAB69751

The mouse homologous cDNAs encoding the polypeptides of the invention were identified as GenBank Accession Numbers XM_132261 (for the mouse homolog of CG14816), XM_127934 (for the mouse homolog of CG30346), NM_008753 and XM_125700 (for the mouse homolog of Oaz1), NM_010952 (for the mouse homolog of Oaz2), NM_016901 (for the mouse homolog of Oaz3), NM_011948 (for the mouse homolog of Mekk1), XM_127816 (for the mouse homolog of PPP2R2A), XM_133882 (for the mouse homolog of PPP2R2D), NM_028392 (for the mouse homolog of PPP2R2B), XM_144326 (for the mouse homolog of PPP2R2C), XM_127113 (for the mouse homolog of Ppp2r5c), and NM_009358 and XM_128662 (for the mouse homolog of Ppp2r5d).

Example 4: dUCPy modifier screen

Expression of Drosophila uncoupling protein dUCPy in a non-vital organ like the eye (Gal4 under control of the eye-specific promoter of the eyeless gene) results in flies with visibly damaged eyes. This easily visible eye phenotype is the basis of a genetic screen for gene products that can modify UCP activity.

Parts of the genomes of the strain with Gal4 expression in the eye and the strain carrying the pUAST-dUCPy construct were combined on one chromosome using genomic recombination. The resulting fly strain has eyes that are permanently damaged by dUCPy expression. Flies of this strain were crossed with flies of a large collection of mutagenized fly strains. In this mutant collection a special expression system (EP-element, Ref.: Rørth P, Proc Natl Acad Sci U S A 1996, 93(22):12418-22) is integrated randomly in different genomic loci. The yeast transcription factor Gal4 can bind to the EP-element and activate the transcription of endogenous genes close the integration site of the EP-element. The activation of the genes therefore occurs in the same cells (eye) that overexpress dUCPy. Since the mutant collection contains several thousand strains with different integration sites of the EP-element it is possible to test a large number of genes whether their expression interacts with dUCPy activity. In case a gene acts as an enhancer of UCP activity the eye defect will be worsened; a suppressor will ameliorate the defect.

Using this screen a gene with enhancing activity was discovered that was found to be the PP2A-B' gene in Drosophila.

Claims

1. A pharmaceutical composition comprising a nucleic acid molecule of the phosphoglycerate mutase-like, NAD kinase-like, ornithine decarboxylase antizyme, mitogen-activated protein kinase kinase kinase, and protein phosphatase 2 regulatory subunit B gene family or a polypeptide encoded thereby or a fragment or a variant of said nucleic acid molecule or said polypeptide or an effector of said nucleic acid molecule or polypeptide together with pharmaceutically acceptable carriers, diluents and/or adjuvants.

2. The composition of claim 1, wherein the nucleic acid molecule is a vertebrate or insect CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' nucleic acid, particulary encoding the human CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologs and/or a nucleic molecule which is complementary thereto or a fragment thereof or a variant thereof.

3. The composition of claim 1 or 2, wherein said nucleic acid molecule
   (a) hybridizes at 50°C in a solution containing 1 x SSC and 0.1% SDS to a nucleic acid molecule as defined in claim 2 and/or a nucleic acid molecule which is complementary thereto;
   (b) it is degenerate with respect to the nucleic acid molecule of (a),
   (c) encodes a polypeptide which is at least 85%, preferably at least 90%, more preferably at least 95%, more preferably at least 98% and up to 99,6% identical to the human CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologs, as defined in claim 2;

(d) differs from the nucleic acid molecule of (a) to (c) by mutation and wherein said mutation causes an alteration, deletion, duplication or premature stop in the encoded polypeptide.

4. The composition of any one of claims 1-3, wherein the nucleic acid molecule is a DNA molecule, particularly a cDNA or a genomic DNA.

5. The composition of any one of claims 1-4, wherein said nucleic acid encodes a polypeptide contributing to regulating the energy homeostasis and/or the metabolism of triglycerides.

6. The composition of any one of claims 1-5, wherein said nucleic acid molecule is a recombinant nucleic acid molecule.

7. The composition of any one of claims 1-6, wherein the nucleic acid molecule is a vector, particularly an expression vector.

8. The composition of any one of claims 1-5, wherein the polypeptide is a recombinant polypeptide.

9. The composition of claim 8, wherein said recombinant polypeptide is a fusion polypeptide.

10. The composition of any one of claims 1-7, wherein said nucleic acid molecule is selected from hybridization probes, primers and anti-sense oligonucleotides.

11. The composition of any one of claims 1-10 which is a diagnostic composition.

12. The composition of any one of claims 1-10 which is a therapeutic composition.

13. The composition of any one of claims 1-12 for the manufacture of an agent for detecting and/or verifying, for the treatment, alleviation and/or prevention of an disorders, including metabolic diseases such as obesity and other body-weight regulation disorders as well as related disorders such as eating disorder, cachexia, diabetes mellitus, hypertension, coronary heart disease, hypercholesterolemia, dyslipidemia, osteoarthritis, gallstones, cancer, e.g. cancers of the reproductive organs, and sleep apnea and others, in cells, cell masses, organs and/or subjects.

14. Use of a nucleic acid molecule of the phosphoglycerate mutase-like, NAD kinase-like, ornithine decarboxylase antizyme, mitogen-activated protein kinase kinase kinase, and protein phosphatase 2 regulatory subunit B gene family or a polypeptide encoded thereby or a fragment or a variant of said nucleic acid molecule or said polypeptide or an effector of said nucleic or polypeptide for controlling the function of a gene and/or a gene product which is influenced and/or modified by a CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologous polypeptide.

15. Use of the nucleic acid molecule of the phosphoglycerate mutase-like, NAD kinase-like, ornithine decarboxylase antizyme, mitogen-activated protein kinase kinase kinase, and protein phosphatase 2 regulatory subunit B gene family or a polypeptide encoded thereby or a fragment or a variant of said nucleic acid molecule or said polypeptide or an effector of said nucleic acid molecule or polypeptide for identifying substances capable of interacting with a CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologous polypeptide.

16. A non-human transgenic animal exhibiting a modified expression of a CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologous polypeptide.

17. The animal of claim 16, wherein the expression of the CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologous polypeptide is increased and/or reduced.

18. A recombinant host cell exhibiting a modified expression of a CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologous polypeptide.

19. The cell of claim 18 which is a human cell.

20. A method of identifying a (poly)peptide involved in the regulation of energy homeostasis and/or metabolism of triglycerides in a mammal comprising the steps of

(a) contacting a collection of (poly)peptides with a CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologous polypeptide or a fragment thereof under conditions that allow binding of said (poly)peptides;

(b) removing (poly)peptides which do not bind and

(c) identifying (poly)peptides that bind to said CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologous polypeptide.

21. A method of screening for an agent which modulates the interaction of a CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologous polypeptide with a binding target/agent, comprising the steps of

(a) incubating a mixture comprising

(aa) a CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologous polypeptide or a fragment thereof;

(ab) a binding target/agent of said CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologous polypeptide or fragment thereof; and

(ac) a candidate agent

under conditions whereby said CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' polypeptide or fragment thereof specifically binds to said binding target/agent at a reference affinity;

(b) detecting the binding affinity of said CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' polypeptide or fragment thereof to said binding target to determine an (candidate) agent-biased affinity; and

(c) determining a difference between (candidate) agent-biased affinity and the reference affinity.

22. A method of producing a composition comprising the (poly)peptide identified by the method of claim 20 or the agent identified by the method of claim 21 with a pharmaceutically acceptable carrier, diluent and/or adjuvant.

23. The method of claim 22 wherein said composition is a pharmaceutical composition for preventing, alleviating or treating of diseases and disorders, including metabolic diseases such as obesity and other body-weight regulation disorders as well as related disorders such as eating disorder, cachexia, diabetes mellitus, hypertension, coronary heart disease, hypercholesterolemia, dyslipidemia, osteoarthritis, gallstones, cancer, e.g. cancers of the reproductive organs, and sleep apnea and other diseases and disorders.

24. Use of a (poly)peptide as identified by the method of claim 20 or of an agent as identified by the method of claim 21 for the preparation

of a pharmaceutical composition for the treatment, alleviation and/or prevention of of diseases and disorders, including metabolic diseases such as obesity and other body-weight regulation disorders as well as related disorders such as eating disorder, cachexia, diabetes mellitus, hypertension, coronary heart disease, hypercholesterolemia, dyslipidemia, osteoarthritis, gallstones, cancer, e.g. cancers of the reproductive organs, and sleep apnea and other diseases and disorders.

25. Use of a nucleic acid molecule of the phosphoglycerate mutase-like, NAD kinase-like, ornithine decarboxylase antizyme, mitogen-activated protein kinase kinase kinase, and protein phosphatase 2 regulatory subunit B gene family or of a fragment thereof for the preparation of a non-human animal which over- or under-expresses the CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' gene product.

26. Kit comprising at least one of

(a) a CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' nucleic acid molecule or a fragment thereof;

(b) a vector comprising the nucleic acid of (a);

(c) a host cell comprising the nucleic acid of (a) or the vector of (b);

(d) a polypeptide encoded by the nucleic acid of (a);

(e) a fusion polypeptide encoded by the nucleic acid of (a);

(f) an antibody, an aptamer or another receptor against the nucleic acid of (a) or the polypeptide of (d) or (e) and

(g) an anti-sense oligonucleotide of the nucleic acid of (a).

Abstract

The present invention discloses CG14816, CG30346, guf, Mekk1, tws, or PP2A-B' homologous proteins regulating the energy homeostasis and the metabolism of triglycerides, and polynucleotides, which identify and encode the proteins disclosed in this invention. The invention also relates to the use of these sequences in the diagnosis, study, prevention, and treatment of diseases and disorders, for example, but not limited to, metabolic disorders and diseases such as the metabolic syndrome, including obesity, as well as related disorders such as eating disorder, cachexia, diabetes mellitus, hypertension, coronary heart disease, hypercholesterolemia, dyslipidemia, osteoarthritis, gallstones, cancers of the reproductive organs, and sleep apnea.

ld 22.10.2002

FIGURE 1. Triglyceride content of a Drosophila CG14816 (GadFly Accession Number) mutant

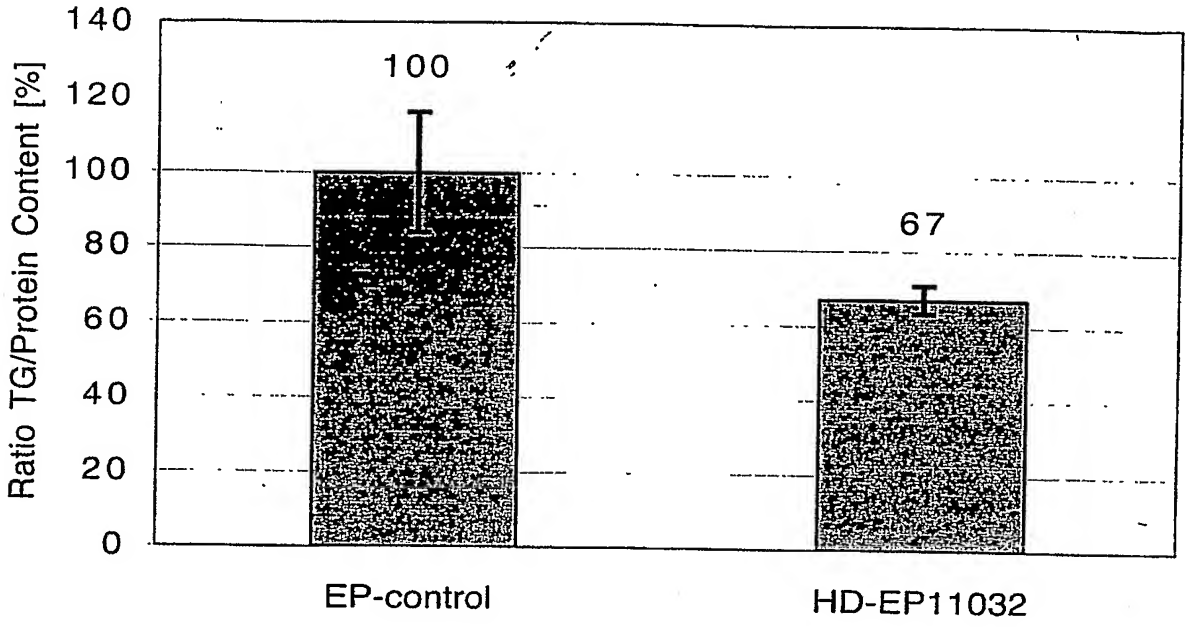

FIGURE 2. Molecular organization of CG14816 gene (GadFly Accession Number)

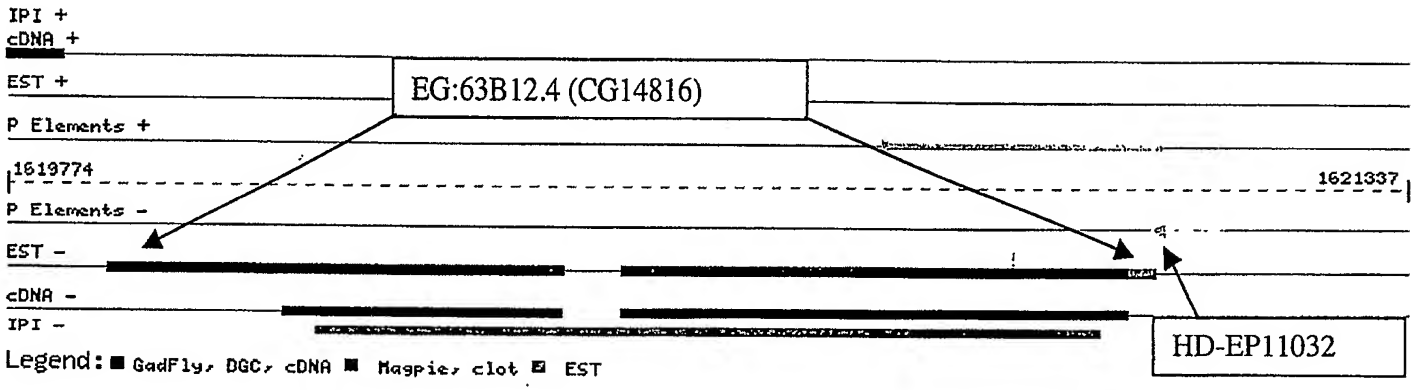

FIGURE 3. Triglyceride content of a Drosophila CG30346 (GadFly Accession Number) mutant

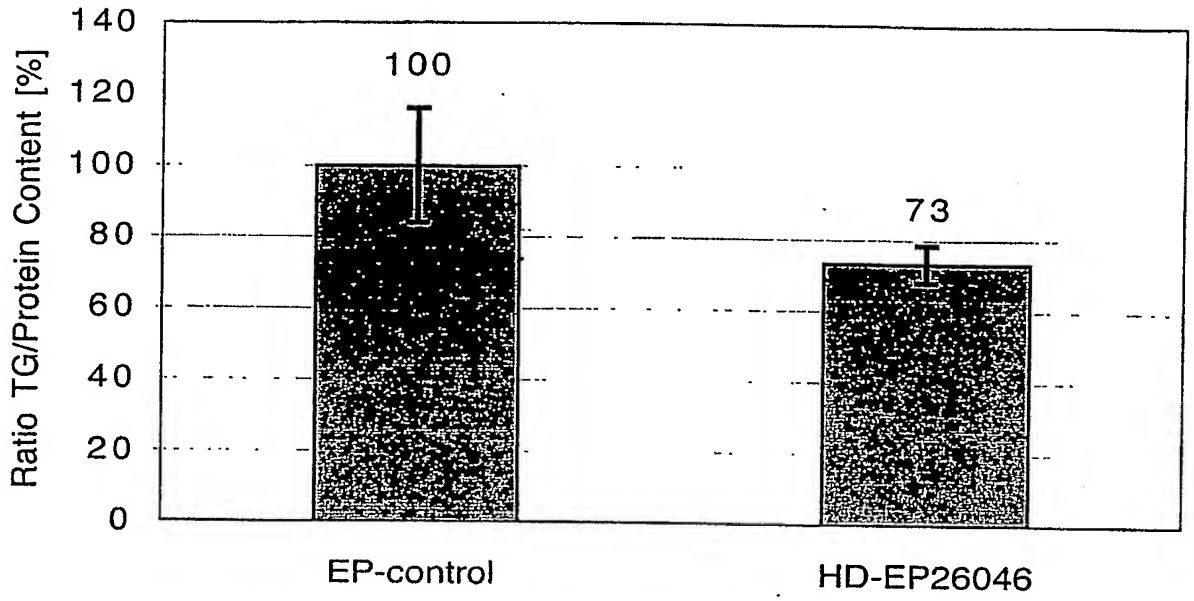

FIGURE 4. Molecular organization of the CG30346 gene (GadFly Accession Number)

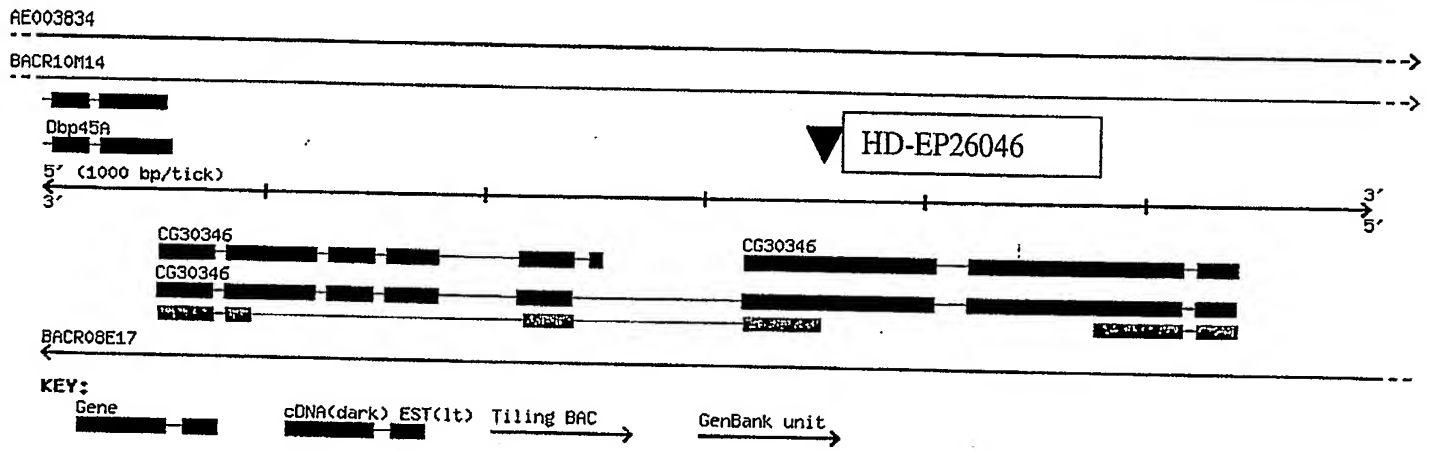

FIGURE 5. Triglyceride content of a Drosophila guf (GadFly Accession Number CG16747) mutant

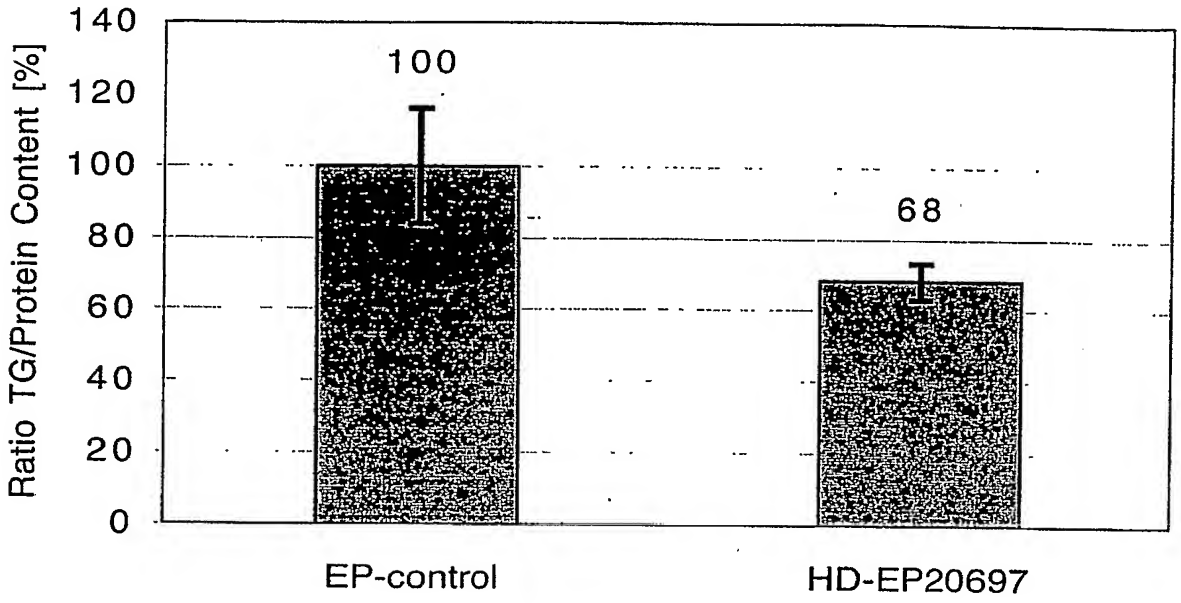

FIGURE 6. Molecular organization of the guf gene (GadFly Accession Number CG16747)

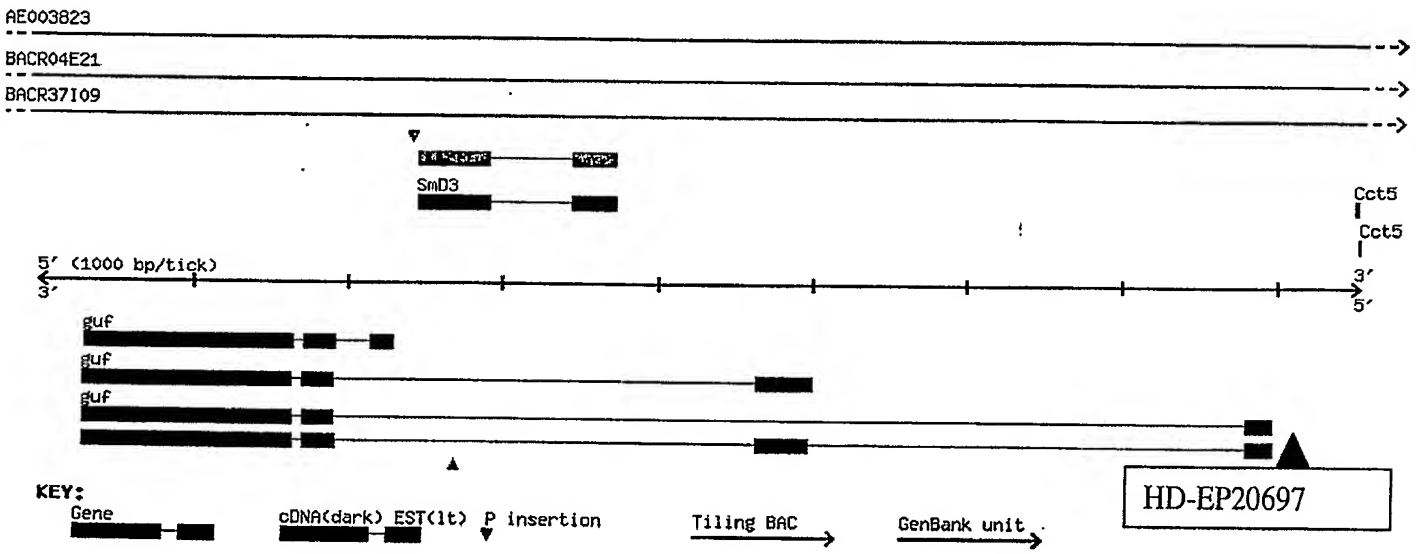

FIGURE 7. Triglyceride content of a Drosophila Mekk1 (GadFly Accession Number CG7717) mutant

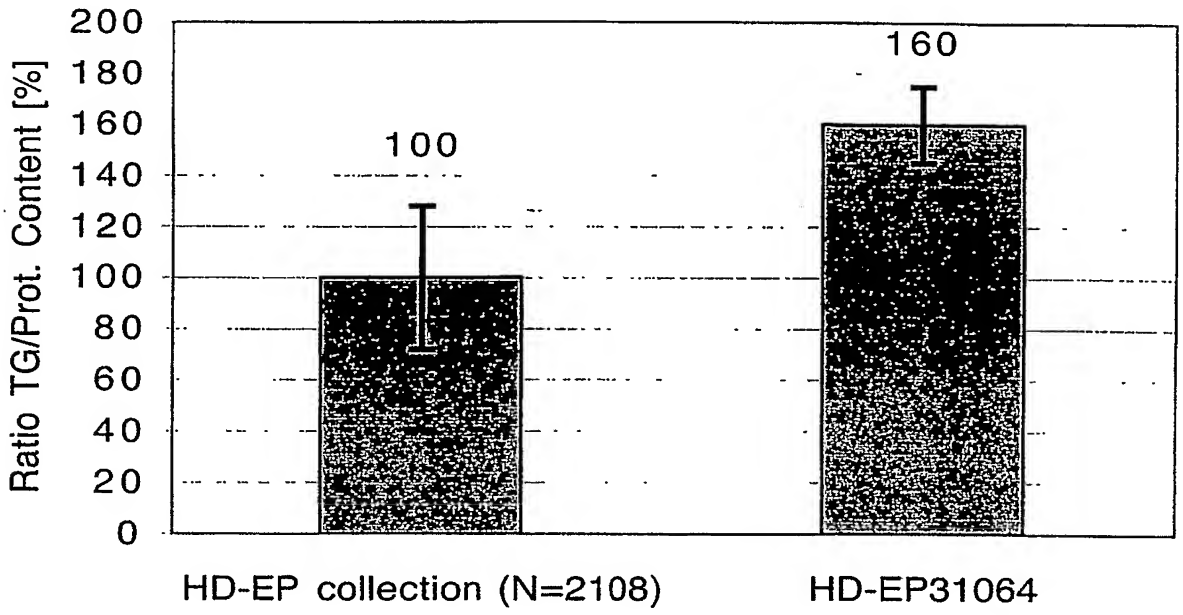

FIGURE 8. Molecular organization of the Mekk1 gene (GadFly Accession Number CG7717)

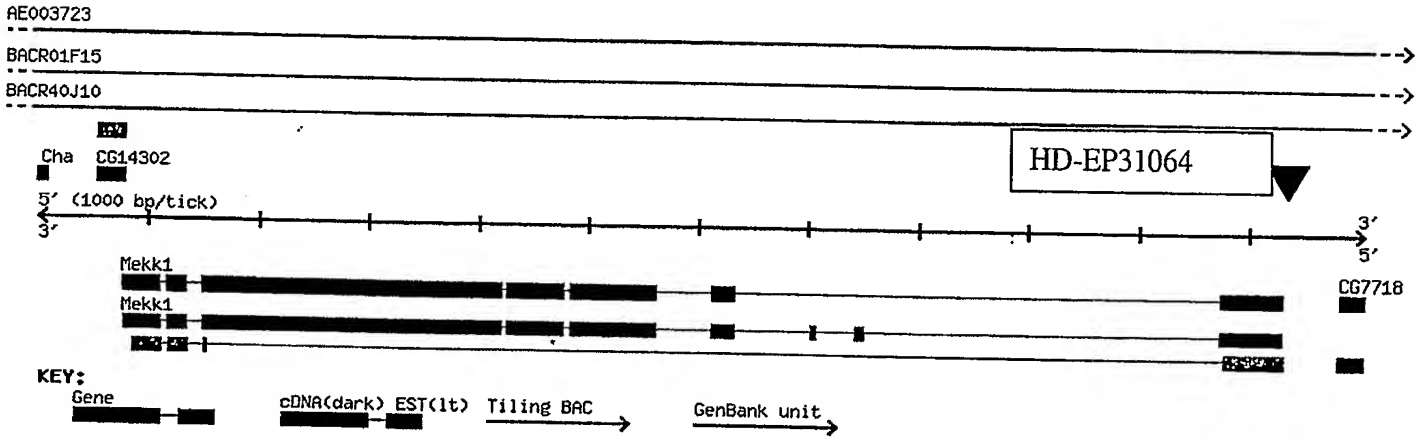

FIGURE 9. Triglyceride content of a Drosophila tws (GadFly Accession Number CG6235) mutant

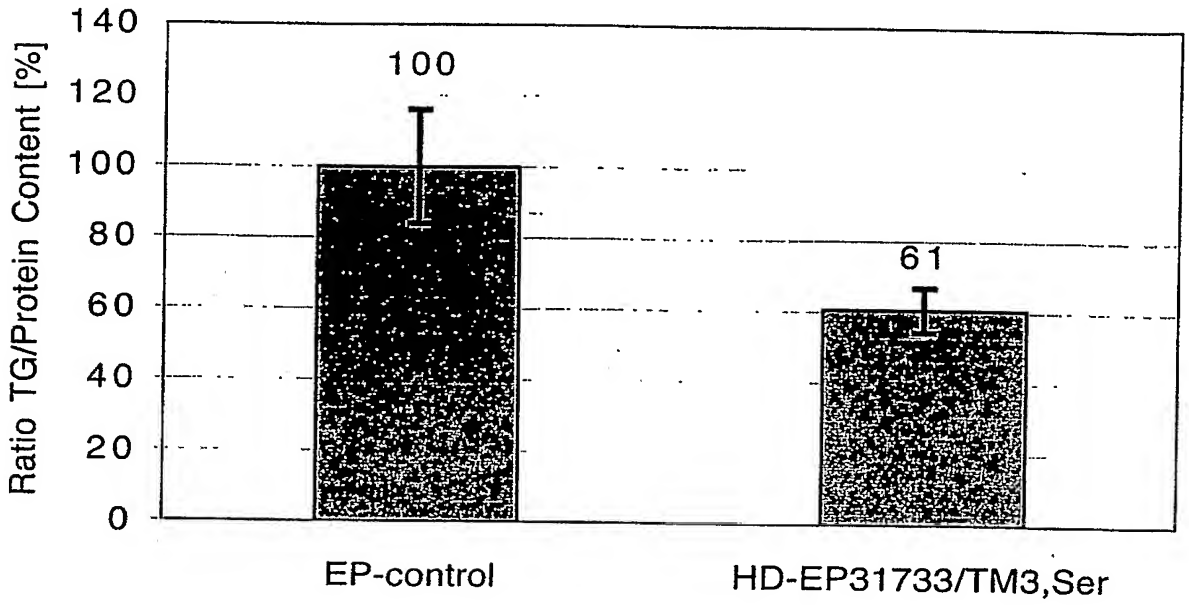

FIGURE 10. Molecular organization of the tws gene (GadFly Accession Number CG6235)

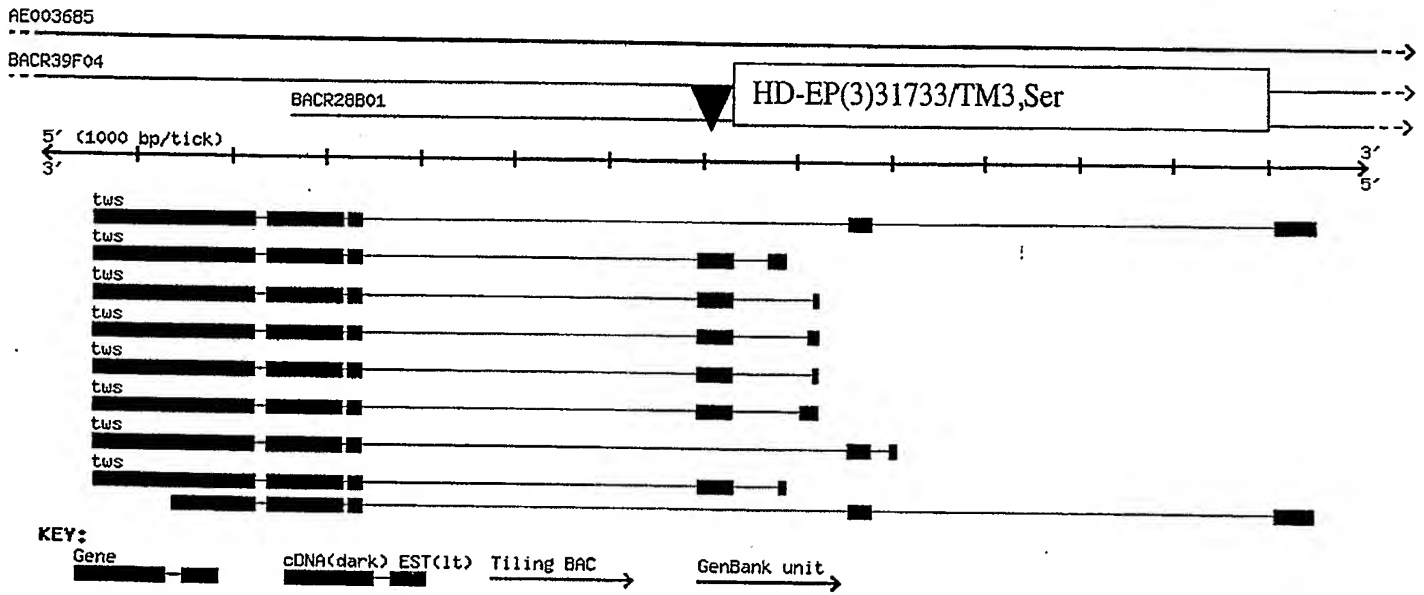

FIGURE 11. Triglyceride content of a Drosophila PP2A-B' (GadFly Accession Number CG7913) mutant

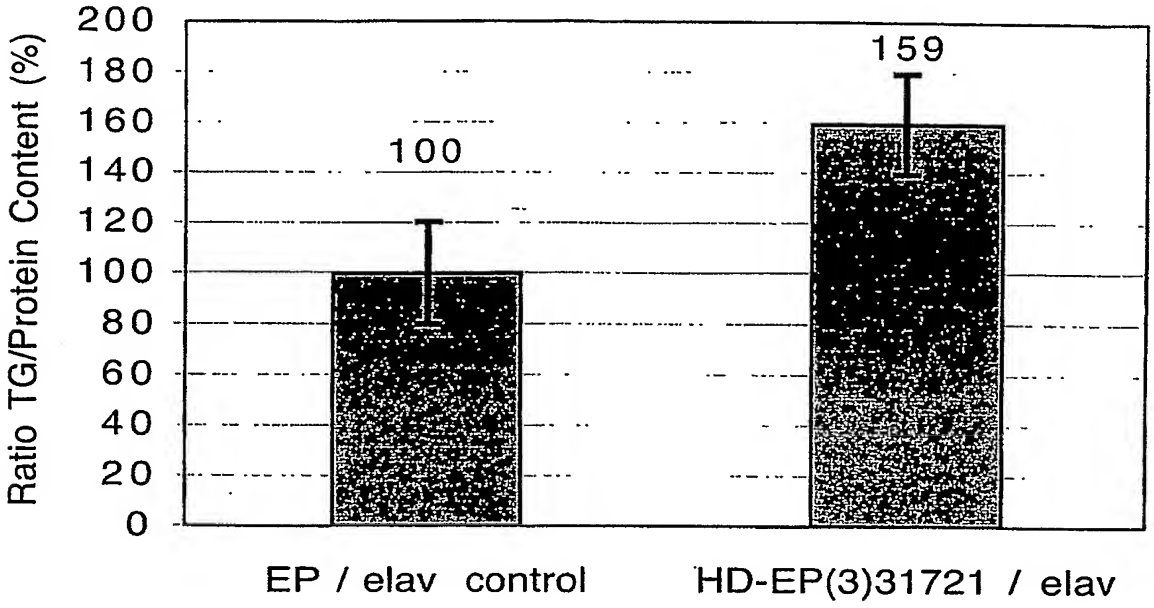

FIGURE 12. Molecular organization of the PP2A-B' gene (GadFly Accession Number CG7913)